Fig. 1A

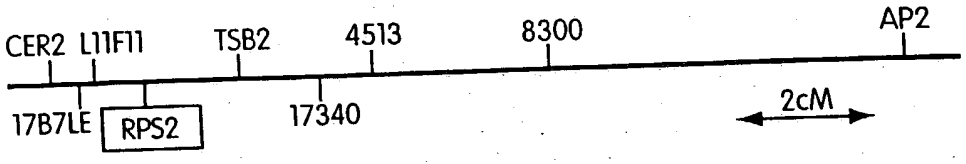


Fig. 1B

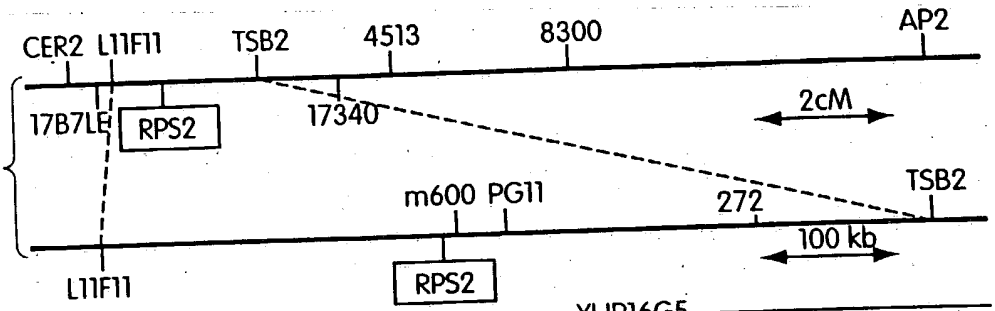


Fig. 1C

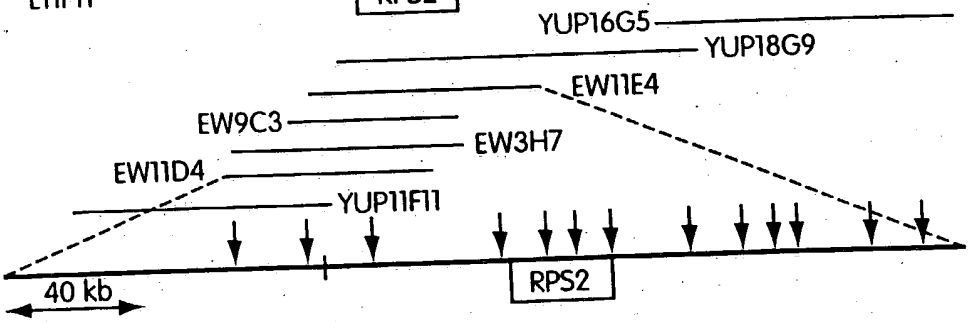


Fig. 1D

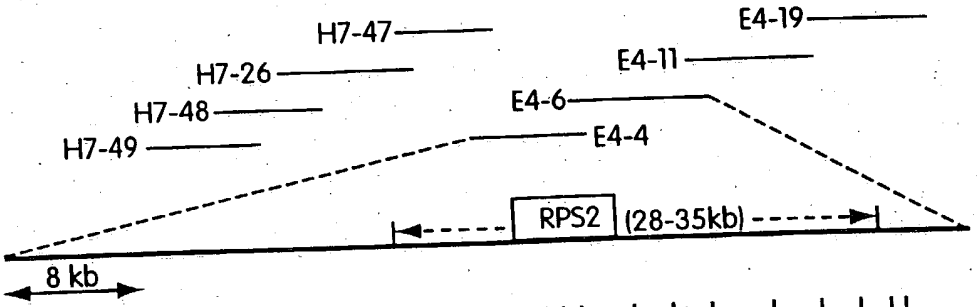


Fig. 1E

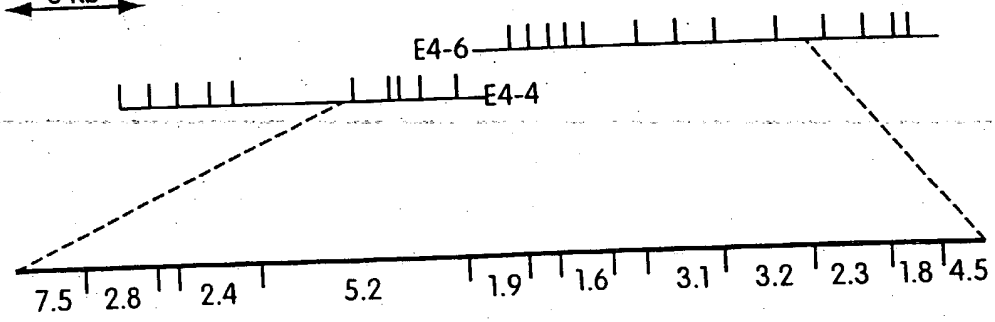
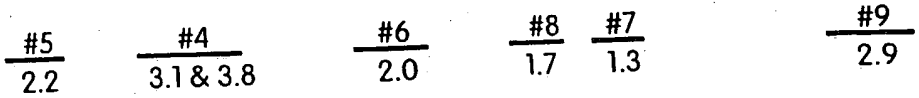


Fig. 1F



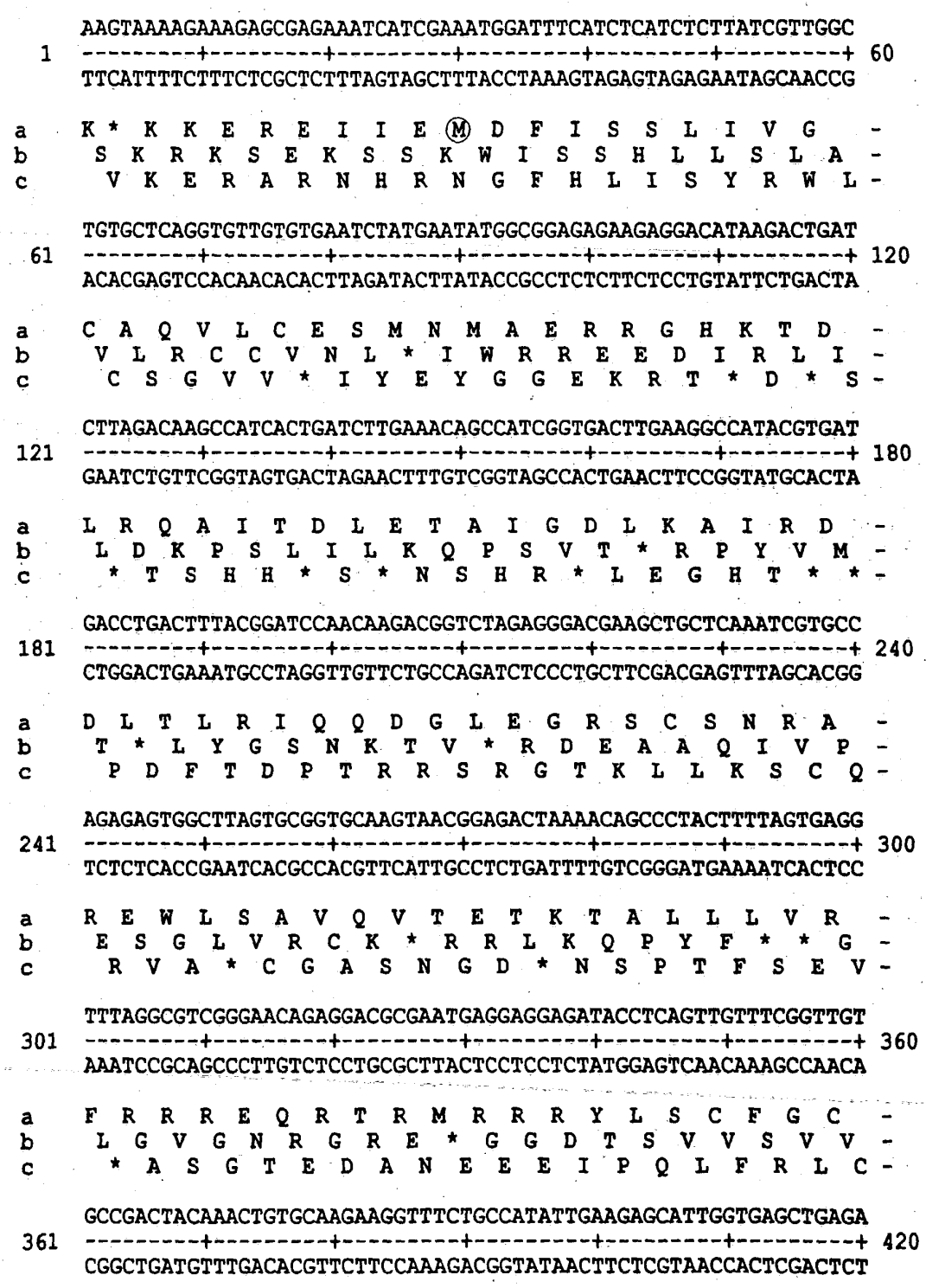

AAGTAAAAGAAAGAGCGAGAAATCATCGAAATGGATTTCATCTCATCTCTTATCGTTGGC
1
60
TTCATTTTCTTTCTCGCTCTTTAGTAGCTTTACCTAAAGTAGAGTAGAGAATAGCAACCG
a K * K K E R E I I E M D F I S S L I V G -
b S K R K S E K S S K W I S S H L L S L A -
c V K E R A R N H R N G F H L I S Y R W L -
TGTGCTCAGGTGTTGTGTGAATCTATGAATATGGCGGAGAGAAGAGGACATAAGACTGAT
61
120
ACACGAGTCCACAACACACTTAGATACTTATACCGCCTCTCTTCTCCTGTATTCTGACTA
a C A Q V L C E S M N M A E R R G H K T D -
b V L R C C V N L * I W R R E E D I R L I -
c C S G V V * I Y E Y G G E K R T * D * S -
CTTAGACAAGCCATCACTGATCTTGAAACAGCCATCGGTGACTTGAAGGCCATACGTGAT
121
180
GAATCTGTTCGGTAGTGACTAGAACTTTGTCGGTAGCCACTGAACTTCCGGTATGCACTA
a L R Q A I T D L E T A I G D L K A I R D -
b L D K P S L I L K Q P S V T * R P Y V M -
c * T S H H * S * N S H R * L E G H T * * -
GACCTGACTTTACGGATCCAACAAGACGGTCTAGAGGGACGAAGCTGCTCAAATCGTGCC
181
240
CTGGACTGAAATGCCTAGGTTGTTCTGCCAGATCTCCCTGCTTCGACGAGTTTAGCACGG
a D L T L R I Q Q D G L E G R S C S N R A -
b T * L Y G S N K T V * R D E A A Q I V P -
c P D F T D P T R R S R G T K L L K S C Q -
AGAGAGTGGCTTAGTGCGGTGCAAGTAACGGAGACTAAAACAGCCCTACTTTTAGTGAGG
241
300
TCTCTCACCGAATCACGCCACGTTCATTGCCTCTGATTTTGTCGGGATGAAAATCACTCC
a R E W L S A V Q V T E T K T A L L L V R -
b E S G L V R C K * R R L K Q P Y F * * G -
c R V A * C G A S N G D * N S P T F S E V -
TTTAGGCGTCGGGAACAGAGGACGCGAATGAGGAGGAGATACCTCAGTTGTTTCGGTTGT
301
360
AAATCCGCAGCCCTTGTCTCCTGCGCTTACTCCTCCTCTATGGAGTCAACAAAGCCAACA
a F R R R E Q R T R M R R R Y L S C F G C -
b L G V G N R G R E * G G D T S V V S V V -
c * A S G T E D A N E E E I P Q L F R L C -
GCCGACTACAAACTGTGCAAGAAGGTTTCTGCCATATTGAAGAGCATTGGTGAGCTGAGA
361
420
CGGCTGATGTTTGACACGTTCTTCCAAAGACGGTATAACTTCTCGTAACCACTCGACTCT

Fig. 2A

```
a    A  D  Y  K  L  C  K  K  V  S  A  I  L  K  S  I  G  E  L  R  -
b     P  T  T  N  C  A  R  R  F  L  P  Y  *  R  A  L  V  S  *  E  -
c      R  L  Q  T  V  Q  E  G  F  C  H  I  E  E  H  W  *  A  E  R -

     GAACGCTCTGAAGCTATCAAAACAGATGGCGGGTCAATTCAAGTAACTTGTAGAGAGATA
421  ---------+---------+---------+---------+---------+---------+ 480
     CTTGCGAGACTTCGATAGTTTTGTCTACCGCCCAGTTAAGTTCATTGAACATCTCTCTAT

a    E  R  S  E  A  I  K  T  D  G  G  S  I  Q  V  T  C  R  E  I  -
b     N  A  L  K  L  S  K  Q  M  A  G  Q  F  K  *  L  V  E  R  Y  -
c      T  L  *  S  Y  Q  N  R  W  R  V  N  S  S  N  L  *  R  D  T -

     CCCATCAAGTCCGTTGTCGGAAATACCACGATGATGGAACAGGTTTTGGAATTTCTCAGT
481  ---------+---------+---------+---------+---------+---------+ 540
     GGGTAGTTCAGGCAACAGCCTTTATGGTGCTACTACCTTGTCCAAAACCTTAAAGAGTCA

a    P  I  K  S  V  V  G  N  T  T  M  M  E  Q  V  L  E  F  L  S  -
b     P  S  S  P  L  S  E  I  P  R  *  W  N  R  F  W  N  F  S  V  -
c      H  Q  V  R  C  R  K  Y  H  D  D  G  T  G  F  G  I  S  Q  * -

     GAAGAAGAAGAAAGAGGAATCATTGGTGTTTATGGACCTGGTGGGGTTGGGAAGACAACG
541  ---------+---------+---------+---------+---------+---------+ 600
     CTTCTTCTTCTTTCTCCTTAGTAACCACAAATACCTGGACCACCCCAACCCTTCTGTTGC

a    E  E  E  E  R  G  I  I  G  V  Y  G  P  G  G  V  G  K  T  T  -
b     K  K  K  K  E  E  S  L  V  F  M  D  L  V  G  L  G  R  Q  R  -
c      R  R  R  K  R  N  H  W  C  L  W  T  W  W  G  W  E  D  N  V -

     TTAATGCAGAGCATTAACAACGAGCTGATCACAAAAGGACATCAGTATGATGTACTGATT
601  ---------+---------+---------+---------+---------+---------+ 660
     AATTACGTCTCGTAATTGTTGCTCGACTAGTGTTTTCCTGTAGTCATACTACATGACTAA

a    L  M  Q  S  I  N  N  E  L  I  T  K  G  H  Q  Y  D  V  L  I  -
b     *  C  R  A  L  T  T  S  *  S  Q  K  D  I  S  M  M  Y  *  F  -
c      N  A  E  H  *  Q  R  A  D  H  K  R  T  S  V  *  C  T  D  L -

     TGGGTTCAAATGTCCAGAGAATTCGGCGAGTGTACAATTCAGCAAGCCGTTGGAGCACGG
661  ---------+---------+---------+---------+---------+---------+ 720
     ACCCAAGTTTACAGGTCTCTTAAGCCGCTCACATGTTAAGTCGTTCGGCAACCTCGTGCC

a    W  V  Q  M  S  R  E  F  G  E  C  T  I  Q  Q  A  V  G  A  R  -
b     G  F  K  C  P  E  N  S  A  S  V  Q  F  S  K  P  L  E  H  G  -
c      G  S  N  V  Q  R  I  R  R  V  Y  N  S  A  S  R  W  S  T  V -

     TTGGGTTTATCTTGGGACGAGAAGGAGACCGGCGAAAACAGAGCTTTGAAGATATACAGA
721  ---------+---------+---------+---------+---------+---------+ 780
     AACCCAAATAGAACCCTGCTCTTCCTCTGGCCGCTTTTGTCTCGAAACTTCTATATGTCT

a    L  G  L  S  W  D  E  K  E  T  G  E  N  R  A  L  K  I  Y  R  -
b     W  V  Y  L  G  T  R  R  R  P  A  K  T  E  L  *  R  Y  T  E  -
c      G  F  I  L  G  R  E  G  D  R  R  K  Q  S  F  E  D  I  Q  S -

     GCTTTGAGACAGAAACGTTTCTTGTTGTTGCTAGATGATGTCTGGGAAGAGATAGACTTG
781  ---------+---------+---------+---------+---------+---------+ 840
     CGAAACTCTGTCTTTGCAAAGAACAACAACGATCTACTACAGACCCTTCTCTATCTGAAC
```

Fig. 2B

```
a    A  L  R  Q  K  R  F  L  L  L  L  D  D  V  W  E  E  I  D  L  -
b     L  *  D  R  N  V  S  C  C  C  *  M  M  S  G  K  R  *  T  W -
c      F  E  T  E  T  F  L  V  V  A  R  *  C  L  G  R  D  R  L  G-

     GAGAAAACTGGAGTTCCTCGACCTGACAGGGAAAACAAATGCAAGGTGATGTTCACGACA
841  ---------+---------+---------+---------+---------+---------+ 900
     CTCTTTTGACCTCAAGGAGCTGGACTGTCCCTTTTGTTTACGTTCCACTACAAGTGCTGT

a    E  K  T  G  V  P  R  P  D  R  E  N  K  C  K  V  M  F  T  T  -
b     R  K  L  E  F  L  D  L  T  G  K  T  N  A  R  *  C  S  R  H -
c      E  N  W  S  S  S  T  *  Q  G  K  Q  M  Q  G  D  V  H  D  T-

     CGGTCTATAGCATTATGCAACAATATGGGTGCGGAATACAAGTTGAGAGTGGAGTTTCTG
901  ---------+---------+---------+---------+---------+---------+ 960
     GCCAGATATCGTAATACGTTGTTATACCCACGCCTTATGTTCAACTCTCACCTCAAAGAC

a    R  S  I  A  L  C  N  N  M  G  A  E  Y  K  L  R  V  E  F  L  -
b     G  L  *  H  Y  A  T  I  W  V  R  N  T  S  *  E  W  S  F  W -
c      V  Y  S  I  M  Q  Q  Y  G  C  G  I  Q  V  E  S  G  V  S  G-

     GAGAAGAAACACGCGTGGGAGCTGTTCTGTAGTAAGGTATGGAGAAAAGATCTTTTAGAG
961  ---------+---------+---------+---------+---------+---------+ 1020
     CTCTTCTTTGTGCGCACCCTCGACAAGACATCATTCCATACCTCTTTTCTAGAAAATCTC

a    E  K  K  H  A  W  E  L  F  C  S  K  V  W  R  K  D  L  L  E  -
b     R  R  N  T  R  G  S  C  S  V  V  R  Y  G  E  K  I  F  *  S -
c      E  E  T  R  V  G  A  V  L  *  *  G  M  E  K  R  S  F  R  V-

     TCATCATCAATTCGCCGGCTCGCGGAGATTATAGTGAGTAAATGTGGAGGATTGCCACTA
1021 ---------+---------+---------+---------+---------+---------+ 1080
     AGTAGTAGTTAAGCGGCCGAGCGCCTCTAATATCACTCATTTACACCTCCTAACGGTGAT

a    S  S  S  I  R  R  L  A  E  I  I  V  S  K  C  G  G  L  P  L  -
b     H  H  Q  F  A  G  S  R  R  L  *  *  V  N  V  E  D  C  H  *  -
c      I  I  N  S  P  A  R  G  D  Y  S  E  *  M  W  R  I  A  T  S-

     GCGTTGATCACTTTAGGAGGAGCCATGGCTCATAGAGAGACAGAAGAAGAGTGGATCCAT
1081 ---------+---------+---------+---------+---------+---------+ 1140
     CGCAACTAGTGAAATCCTCCTCGGTACCGAGTATCTCTCTGTCTTCTTCTCACCTAGGTA

a    A  L  I  T  L  G  G  A  M  A  H  R  E  T  E  E  E  W  I  H  -
b     R  *  S  L  *  E  E  P  W  L  I  E  R  Q  K  K  S  G  S  M -
c      V  D  H  F  R  R  S  H  G  S  *  R  D  R  R  R  V  D  P  C-

     GCTAGTGAAGTTCTGACTAGATTTCCAGCAGAGATGAAGGGTATGAACTATGTATTTGCC
1141 ---------+---------+---------+---------+---------+---------+ 1200
     CGATCACTTCAAGACTGATCTAAAGGTCGTCTCTACTTCCCATACTTGATACATAAACGG

a    A  S  E  V  L  T  R  F  P  A  E  M  K  G  M  N  Y  V  F  A  -
b     L  V  K  F  *  L  D  F  Q  Q  R  *  R  V  *  T  M  Y  L  P -
c      *  *  S  S  D  *  I  S  S  R  D  E  G  Y  E  L  C  I  C  P-

     CTTTTGAAATTCAGCTACGACAACCTCGAGAGTGATCTGCTTCGGTCTTGTTTCTTGTAC
1201 ---------+---------+---------+---------+---------+---------+ 1260
     GAAAACTTTAAGTCGATGCTGTTGGAGCTCTCACTAGACGAAGCCAGAACAAAGAACATG
```

Fig. 2C

```
a    L  L  K  F  S  Y  D  N  L  E  S  D  L  L  R  S  C  F  L  Y  -
b     F  *  N  S  A  T  T  T  S  R  V  I  C  F  G  L  V  S  C  T  -
c      F  E  I  Q  L  R  Q  P  R  E  *  S  A  S  V  L  F  L  V  L -

     TGCGCTTTATTCCCAGAAGAACATTCTATAGAGATCGAGCAGCTTGTTGAGTACTGGGTC
1261 ---------+---------+---------+---------+---------+---------+ 1320
     ACGCGAAATAAGGGTCTTCTTGTAAGATATCTCTAGCTCGTCGAACAACTCATGACCCAG

a    C  A  L  F  P  E  E  H  S  I  E  I  E  Q  L  V  E  Y  W  V  -
b     A  L  Y  S  Q  K  N  I  L  *  R  S  S  S  L  L  S  T  G  S  -
c      R  F  I  P  R  R  T  F  Y  R  D  R  A  A  C  *  V  L  G  R -

     GGCGAAGGGTTTCTCACCAGCTCCCATGGCGTTAACACCATTTACAAGGGATATTTTCTC
1321 ---------+---------+---------+---------+---------+---------+ 1380
     CCGCTTCCCAAAGAGTGGTCGAGGGTACCGCAATTGTGGTAAATGTTCCCTATAAAAGAG

a    G  E  G  F  L  T  S  S  H  G  V  N  T  I  Y  K  G  Y  F  L  -
b     A  K  G  F  S  P  A  P  M  A  L  T  P  F  T  R  D  I  F  S  -
c      R  R  V  S  H  Q  L  P  W  R  *  H  H  L  Q  G  I  F  S  H -

     ATTGGGGATCTGAAAGCGGCATGTTTGTTGGAAACCGGAGATGAGAAAACACAGGTGAAG
1381 ---------+---------+---------+---------+---------+---------+ 1440
     TAACCCCTAGACTTTCGCCGTACAAACAACCTTTGGCCTCTACTCTTTTGTGTCCACTTC

a    I  G  D  L  K  A  A  C  L  L  E  T  G  D  E  K  T  Q  V  K  -
b     L  G  I  *  K  R  H  V  C  W  K  P  E  M  R  K  H  R  *  R  -
c      W  G  S  E  S  G  M  F  V  G  N  R  R  *  E  N  T  G  E  D -

     ATGCATAATGTGGTCAGAAGCTTTGCATTGTGGATGGCATCTGAACAGGGGACTTATAAG
1441 ---------+---------+---------+---------+---------+---------+ 1500
     TACGTATTACACCAGTCTTCGAAACGTAACACCTACCGTAGACTTGTCCCCTGAATATTC

a    M  H  N  V  V  R  S  F  A  L  W  M  A  S  E  Q  G  T  Y  K  -
b     C  I  M  W  S  E  A  L  H  C  G  W  H  L  N  R  G  L  I  R  -
c      A  *  C  G  Q  K  L  C  I  V  D  G  I  *  T  G  D  L  *  G -

     GAGCTGATCCTAGTTGAGCCTAGCATGGGACATACTGAAGCTCCTAAAGCAGAAAACTGG
1501 ---------+---------+---------+---------+---------+---------+ 1560
     CTCGACTAGGATCAACTCGGATCGTACCCTGTATGACTTCGAGGATTTCGTCTTTTGACC

a    E  L  I  L  V  E  P  S  M  G  H  T  E  A  P  K  A  E  N  W  -
b     S  *  S  *  L  S  L  A  W  D  I  L  K  L  L  K  Q  K  T  G  -
c      A  D  P  S  *  A  *  H  G  T  Y  *  S  S  *  S  R  K  L  A -

     CGACAAGCGTTGGTGATCTCATTGTTAGATAACAGAATCCAGACCTTGCCTGAAAAACTC
1561 ---------+---------+---------+---------+---------+---------+ 1620
     GCTGTTCGCAACCACTAGAGTAACAATCTATTGTCTTAGGTCTGGAACGGACTTTTTGAG

a    R  Q  A  L  V  I  S  L  L  D  N  R  I  Q  T  L  P  E  K  L  -
b     D  K  R  W  *  S  H  C  *  I  T  E  S  R  P  C  L  K  N  S  -
c      T  S  V  G  D  L  I  V  R  *  Q  N  P  D  L  A  *  K  T  H -
```

Fig. 2D

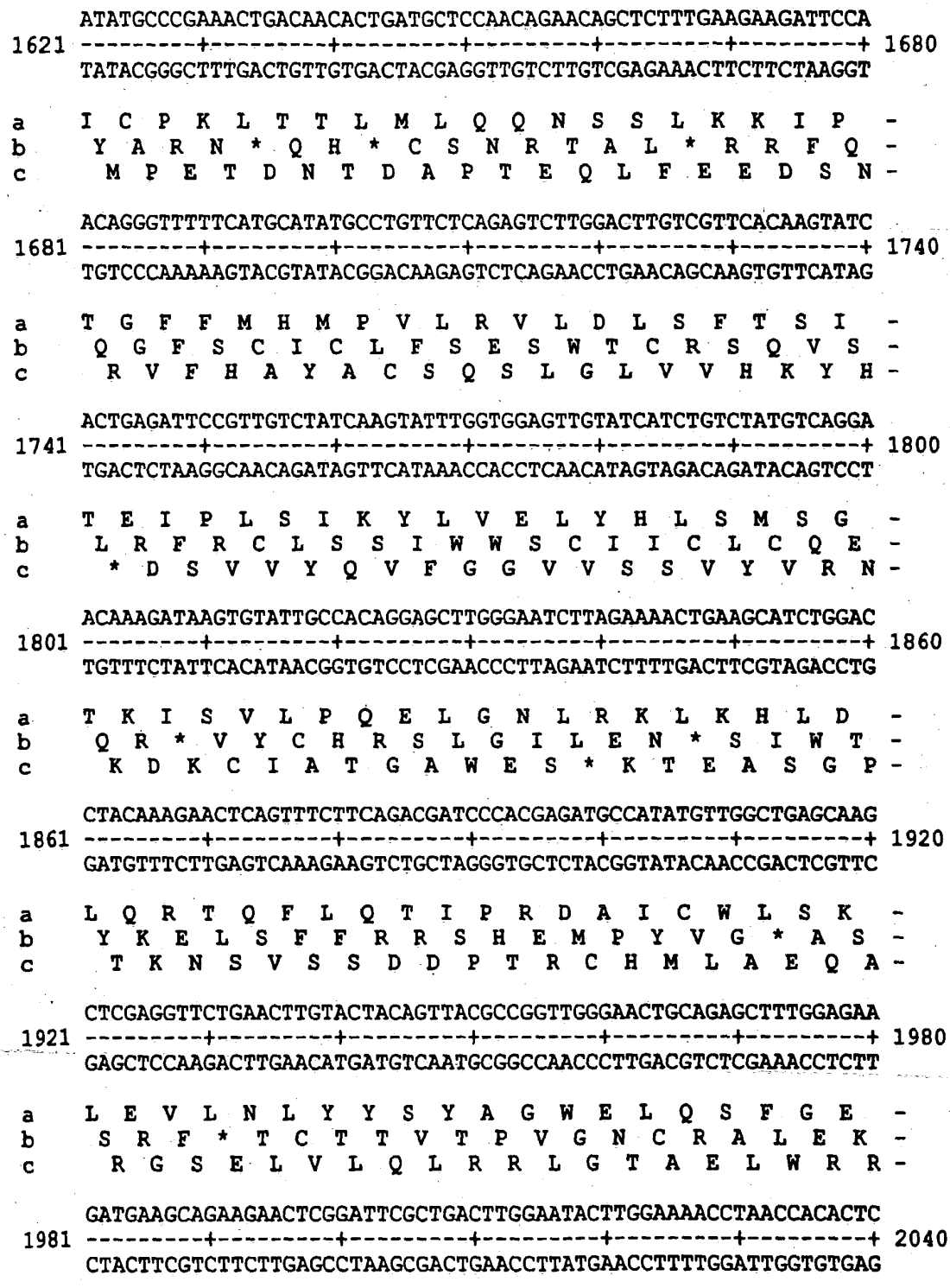

ATATGCCCGAAACTGACAACACTGATGCTCCAACAGAACAGCTCTTTGAAGAAGATTCCA
1621 1680
TATACGGGCTTTGACTGTTGTGACTACGAGGTTGTCTTGTCGAGAAACTTCTTCTAAGGT
a I C P K L T T L M L Q Q N S S L K K I P -
b Y A R N * Q H * C S N R T A L * R R F Q -
c M P E T D N T D A P T E Q L F E E D S N -
ACAGGGTTTTTCATGCATATGCCTGTTCTCAGAGTCTTGGACTTGTCGTTCACAAGTATC
1681 1740
TGTCCCAAAAAGTACGTATACGGACAAGAGTCTCAGAACCTGAACAGCAAGTGTTCATAG
a T G F F M H M P V L R V L D L S F T S I -
b Q G F S C I C L F S E S W T C R S Q V S -
c R V F H A Y A C S Q S L G L V V H K Y H -
ACTGAGATTCCGTTGTCTATCAAGTATTTGGTGGAGTTGTATCATCTGTCTATGTCAGGA
1741 1800
TGACTCTAAGGCAACAGATAGTTCATAAACCACCTCAACATAGTAGACAGATACAGTCCT
a T E I P L S I K Y L V E L Y H L S M S G -
b L R F R C L S S I W W S C I I C L C Q E -
c * D S V V Y Q V F G G V V S S V Y V R N -
ACAAAGATAAGTGTATTGCCACAGGAGCTTGGGAATCTTAGAAAACTGAAGCATCTGGAC
1801 1860
TGTTTCTATTCACATAACGGTGTCCTCGAACCCTTAGAATCTTTTGACTTCGTAGACCTG
a T K I S V L P Q E L G N L R K L K H L D -
b Q R * V Y C H R S L G I L E N * S I W T -
c K D K C I A T G A W E S * K T E A S G P -
CTACAAAGAACTCAGTTTCTTCAGACGATCCCACGAGATGCCATATGTTGGCTGAGCAAG
1861 1920
GATGTTTCTTGAGTCAAAGAAGTCTGCTAGGGTGCTCTACGGTATACAACCGACTCGTTC
a L Q R T Q F L Q T I P R D A I C W L S K -
b Y K E L S F F R R S H E M P Y V G * A S -
c T K N S V S S D D P T R C H M L A E Q A -
CTCGAGGTTCTGAACTTGTACTACAGTTACGCCGGTTGGGAACTGCAGAGCTTTGGAGAA
1921 1980
GAGCTCCAAGACTTGAACATGATGTCAATGCGGCCAACCCTTGACGTCTCGAAACCTCTT
a L E V L N L Y Y S Y A G W E L Q S F G E -
b S R F * T C T T V T P V G N C R A L E K -
c R G S E L V L Q L R R L G T A E L W R R -
GATGAAGCAGAAGAACTCGGATTCGCTGACTTGGAATACTTGGAAAACCTAACCACACTC
1981 2040
CTACTTCGTCTTCTTGAGCCTAAGCGACTGAACCTTATGAACCTTTTGGATTGGTGTGAG

Fig. 2E

```
a    D  E  A  E  E  L  G  F  A  D  L  E  Y  L  E  N  L  T  T  L  -
b     M  K  Q  K  N  S  D  S  L  T  W  N  T  W  K  T  *  P  H  S  -
c      *  S  R  R  T  R  I  R  *  L  G  I  L  G  K  P  N  H  T  R -

     GGTATCACTGTTCTCTCATTGGAGACCCTAAAAACTCTCTTCGAGTTCGGTGCTTTGCAT
2041 ---------+---------+---------+---------+---------+---------+ 2100
     CCATAGTGACAAGAGAGTAACCTCTGGGATTTTTGAGAGAAGCTCAAGCCACGAAACGTA

a    G  I  T  V  L  S  L  E  T  L  K  T  L  F  E  F  G  A  L  H  -
b     V  S  L  F  S  H  W  R  P  *  K  L  S  S  S  S  V  L  C  I  -
c      Y  H  C  S  L  I  G  D  P  K  N  S  L  R  V  R  C  F  A  * -

     AAACATATACAGCATCTCCACGTTGAAGAGTGCAATGAACTCCTCTACTTCAATCTCCCA
2101 ---------+---------+---------+---------+---------+---------+ 2160
     TTTGTATATGTCGTAGAGGTGCAACTTCTCACGTTACTTGAGGAGATGAAGTTAGAGGGT

a    K  H  I  Q  H  L  H  V  E  E  C  N  E  L  L  Y  F  N  L  P  -
b     N  I  Y  S  I  S  T  L  K  S  A  M  N  S  S  T  S  I  S  H  -
c      T  Y  T  A  S  P  R  *  R  V  Q  *  T  P  L  L  Q  S  P  I -

     TCACTCACTAACCATGGCAGGAACCTGAGAAGACTTAGCATTAAAAGTTGCCATGACTTG
2161 ---------+---------+---------+---------+---------+---------+ 2220
     AGTGAGTGATTGGTACCGTCCTTGGACTCTTCTGAATCGTAATTTTCAACGGTACTGAAC

a    S  L  T  N  H  G  R  N  L  R  R  L  S  I  K  S  C  H  D  L  -
b     H  S  L  T  M  A  G  T  *  E  D  L  A  L  K  V  A  M  T  W  -
c      T  H  *  P  W  Q  E  P  E  K  T  *  H  *  K  L  P  *  L  G -

     GAGTACCTGGTCACACCCGCAGATTTTGAAAATGATTGGCTTCCGAGTCTAGAGGTTCTG
2221 ---------+---------+---------+---------+---------+---------+ 2280
     CTCATGGACCAGTGTGGGCGTCTAAAACTTTTACTAACCGAAGGCTCAGATCTCCAAGAC

a    E  Y  L  V  T  P  A  D  F  E  N  D  W  L  P  S  L  E  V  L  -
b     S  T  W  S  H  P  Q  I  L  K  M  I  G  F  R  V  *  R  F  *  -
c      V  P  G  H  T  R  R  F  *  K  *  L  A  S  E  S  R  G  S  D -

     ACGTTACACAGCCTTCACAACTTAACCAGAGTGTGGGGAAATTCTGTAAGCCAAGATTGT
2281 ---------+---------+---------+---------+---------+---------+ 2340
     TGCAATGTGTCGGAAGTGTTGAATTGGTCTCACACCCCTTTAAGACATTCGGTTCTAACA

a    T  L  H  S  L  H  N  L  T  R  V  W  G  N  S  V  S  Q  D  C  -
b     R  Y  T  A  F  T  T  *  P  E  C  G  E  I  L  *  A  K  I  V  -
c      V  T  Q  P  S  Q  L  N  Q  S  V  G  K  F  C  K  P  R  L  S -

     CTGCGGAATATCCGTTGCATAAACATTTCACACTGCAACAAGCTGAAGAATGTCTCATGG
2341 ---------+---------+---------+---------+---------+---------+ 2400
     GACGCCTTATAGGCAACGTATTTGTAAAGTGTGACGTTGTTCGACTTCTTACAGAGTACC

a    L  R  N  I  R  C  I  N  I  S  H  C  N  K  L  K  N  V  S  W  -
b     C  G  I  S  V  A  *  T  F  H  T  A  T  S  *  R  M  S  H  G  -
c      A  E  Y  P  L  H  K  H  F  T  L  Q  Q  A  E  E  C  L  M  G -

     GTTCAGAAACTCCCAAAGCTAGAGGTGATTGAACTGTTCGACTGCAGAGAGATAGAGGAA
2401 ---------+---------+---------+---------+---------+---------+ 2460
     CAAGTCTTTGAGGGTTTCGATCTCCACTAACTTGACAAGCTGACGTCTCTCTATCTCCTT
```

Fig. 2F

```
a    V  Q  K  L  P  K  L  E  V  I  E  L  F  D  C  R  E  I  E  E  -
b     F  R  N  S  Q  S  *  R  *  L  N  C  S  T  A  E  R  *  R  N  -
c      S  E  T  P  K  A  R  G  D  *  T  V  R  L  Q  R  D  R  G  I  -

     TTGATAAGCGAACACGAGAGTCCATCCGTCGAAGATCCAACATTGTTCCCAAGCCTGAAG
2461 ---------+---------+---------+---------+---------+---------+ 2520
     AACTATTCGCTTGTGCTCTCAGGTAGGCAGCTTCTAGGTTGTAACAAGGGTTCGGACTTC

a    L  I  S  E  H  E  S  P  S  V  E  D  P  T  L  F  P  S  L  K  -
b     *  *  A  N  T  R  V  H  P  S  K  I  Q  H  C  S  Q  A  *  R  -
c      D  K  R  T  R  E  S  I  R  R  R  S  N  I  V  P  K  P  E  D  -

     ACCTTGAGAACTAGGGATCTGCCAGAACTAAACAGCATCCTCCCATCTCGATTTTCATTC
2521 ---------+---------+---------+---------+---------+---------+ 2580
     TGGAACTCTTGATCCCTAGACGGTCTTGATTTGTCGTAGGAGGGTAGAGCTAAAAGTAAG

a    T  L  R  T  R  D  L  P  E  L  N  S  I  L  P  S  R  F  S  F  -
b     P  *  E  L  G  I  C  Q  N  *  T  A  S  S  H  L  D  F  H  S  -
c      L  E  N  *  G  S  A  R  T  K  Q  H  P  P  I  S  I  F  I  P  -

     CAAAAAGTTGAAACATTAGTCATCACAAATTGCCCCAGAGTTAAGAAACTGCCGTTTCAG
2581 ---------+---------+---------+---------+---------+---------+ 2640
     GTTTTTCAACTTTGTAATCAGTAGTGTTTAACGGGGTCTCAATTCTTTGACGGCAAAGTC

a    Q  K  V  E  T  L  V  I  T  N  C  P  R  V  K  K  L  P  F  Q  -
b     K  K  L  K  H  *  S  S  Q  I  A  P  E  L  R  N  C  R  F  R  -
c      K  S  *  N  I  S  H  H  K  L  P  Q  S  *  E  T  A  V  S  G  -

     GAGAGGAGGACCCAGATGAACTTGCCAACAGTTTATTGTGAGGAGAAATGGTGGAAAGCA
2641 ---------+---------+---------+---------+---------+---------+ 2700
     CTCTCCTCCTGGGTCTACTTGAACGGTTGTCAAATAACACTCCTCTTTACCACCTTTCGT

a    E  R  R  T  Q  M  N  L  P  T  V  Y  C  E  E  K  W  W  K  A  -
b     R  G  G  P  R  *  T  C  Q  Q  F  I  V  R  R  N  G  G  K  H  -
c      E  E  D  P  D  E  L  A  N  S  L  L  *  G  E  M  V  E  S  T  -

     CTGGAAAAAGATCAACCAAACGAAGAGCTTTGTTATTTACCGCGCTTTGTTCCAAATTGA
2701 ---------+---------+---------+---------+---------+---------+ 2760
     GACCTTTTTCTAGTTGGTTTGCTTCTCGAAACAATAAATGGCGCGAAACAAGGTTTAACT

a    L  E  K  D  Q  P  N  E  E  L  C  Y  L  P  R  F  V  P  N  *  -
b     W  K  K  I  N  Q  T  K  S  F  V  I  Y  R  A  L  F  Q  I  D  -
c      G  K  R  S  T  K  R  R  A  L  L  F  T  A  L  C  S  K  L  I  -

     TATAAGAGCTAAGAGCACTCTGTACAAATATGTCCATTCATAAGATGCAGGAAGCCAGGA
2761 ---------+---------+---------+---------+---------+---------+ 2820
     ATATTCTCGATTCTCGTGAGACATGTTTATACAGGTAAGTATTCTACGTCCTTCGGTCCT

a    Y  K  S  *  E  H  S  V  Q  I  C  P  F  I  R  C  R  K  P  G  -
b     I  R  A  K  S  T  L  Y  K  Y  V  H  S  *  D  A  G  S  Q  E  -
c      *  E  L  R  A  L  C  T  N  M  S  I  H  K  M  Q  E  A  R  K  -

     AGGTTGTTCCAGTGAAGTCATCAACTTTCCACATAGCCACAAAACTAGAGATTATGTAAT
2821 ---------+---------+---------+---------+---------+---------+ 2880
     TCCAACAAGGTCACTTCAGTAGTTGAAAGGTGTATCGGTGTTTTGATCTCTAATACATTA
```

Fig. 2G

```
a    R  L  F  Q  *  S  H  Q  L  S  T  *  P  Q  N  *  R  L  C  N  -
b     G  C  S  S  E  V  I  N  F  P  H  S  H  K  T  R  D  Y  V  I  -
c      V  V  P  V  K  S  S  T  F  H  I  A  T  K  L  E  I  M  *  S  -

     CATAAAAACCAAACTATCCGCGA
2881 ---------+---------+--- 2903
     GTATTTTTGGTTTGATAGGCGCT

a    H  K  N  Q  T  I  R     -
b     I  K  T  K  L  S  A     -
c      *  K  P  N  Y  P  R     -

 ENZYMES THAT DO CUT:

      NONE

 ENZYMES THAT DO NOT CUT:

      KpnI
```

Fig. 2H

```
-146
ATCGATTGATCTCTGGCTCAGTGCGAGTAGTCCATTTGAGAGCAGTCGTAGCCCCGCGTG  -86

GCGCATCATGGAGCTATTTGGAATTTTCGCAGGGTTATCGATTCGTAGTGGGAACCCATT  -26

                    1
CATTGTTTGGAACCACCAACGGACGACTTAACAAGCTCCCCGAGGTGCATGATGAAAATT   35
                                                 MetLysIle

GCTCCAGTTGCCATAAATCACAGCCCGCTCAGCAGGGAGGTCCCGTCACACGCGGCACCC   95
AlaProValAlaIleAsnHisSerProLeuSerArgGluValProSerHisAlaAlaPro

ACTCAGGCAAAGCAAACCAACCTTCAATCTGAAGCTGGCGATTTAGATGCAAGAAAAAGT  155
ThrGlnAlaLysGlnThrAsnLeuGlnSerGluAlaGlyAspLeuAspAlaArgLysSer

AGCGCTTCAAGCCCGGAAACCCGCGCATTACTCGCTACTAAGACAGTACTCGGGAGACAC  215
SerAlaSerSerProGluThrArgAlaLeuLeuAlaThrLysThrValLeuGlyArgHis

AAGATAGAGGTTCCGGCCTTTGGAGGGTGGTTCAAAAAGAAATCATCTAAGCACGAGACG  275
LysIleGluValProAlaPheGlyGlyTrpPheLysLysLysSerSerLysHisGluThr

GGCGGTTCAAGTGCCAACGCAGATAGTTCGAGCGTGGCTTCCGATTCCACCGAAAAACCT  335
GlyGlySerSerAlaAsnAlaAspSerSerSerValAlaSerAspSerThrGluLysPro

TTGTTCCGTCTCACGCACGTTCCTTACGTATCCCAAGGTAATGAGCGAATGGGATGTTGG  395
LeuPheArgLeuThrHisValProTyrValSerGlnGlyAsnGluArgMetGlyCysTrp

TATGCCTGCGCAAGAATGGTTGGCCATTCTGTCGAAGCTGGGCCTCGCCTAGGGCTGCCG  455
TyrAlaCysAlaArgMetValGlyHisSerValGluAlaGlyProArgLeuGlyLeuPro

GAGCTCTATGAGGGAAGGGAGGCGCCAGCTGGGCTACAAGATTTTTCAGATGTAGAAAGG  515
GluLeuTyrGluGlyArgGluAlaProAlaGlyLeuGlnAspPheSerAspValGluArg

TTTATTCACAATGAAGGATTAACTCGGGTAGACCTTCCAGACAATGAGAGATTTACACAC  575
PheIleHisAsnGluGlyLeuThrArgValAspLeuProAspAsnGluArgPheThrHis
```

Fig. 3A

```
GAAGAGTTGGGTGCACTGTTGTATAAGCACGGGCCGATTATATTTGGGTGGAAAACTCCG  635
GluGluLeuGlyAlaLeuLeuTyrLysHisGlyProIleIlePheGlyTrpLysThrPro

AATGACAGCTGGCACATGTCGGTCCTCACTGGTGTCGATAAAGAGACGTCGTCCATTACT  695
AsnAspSerTrpHisMetSerValLeuThrGlyValAspLysGluThrSerSerIleThr

TTTCACGATCCCCGACAGGGGCCGGACCTAGCAATGCCGCTCGATTACTTTAATCAGCGA  755
PheHisAspProArgGlnGlyProAspLeuAlaMetProLeuAspTyrPheAsnGlnArg

TTGGCATGGCAGGTTCCACACGCAATGCTCTACCGCTAAGTAGCAGGGTATCTTCACGTG  815
LeuAlaTrpGlnValProHisAlaMetLeuTyrArgEnd

GCGGCATCATGACAAGCCCATGATGCCGCCAGCAGCTACCTGAATGCCGTCTGGCTTTTT  875
                ---------->          <----------

GGTCCCTATTGTCGTATCCGGAAGATGACGTCAAAGAATCTCGGCAAGAGCTTTCTTGCT  935

CGACTCCTCAGCTTCCGGATCGATCAGGTCGCTTGCCAGAGCGCGCTTGTCCATGAGCAT  995

CTGCCACAGCTGCTGGTCGATGGTGTCCTCAGCTAAAGGGATTTTGACGACAACCATGCG 1055

CAACTGCCCGTTGCGATACGCTCGATCCTGAAGCCCCGGTGTCCATGGCAGCCCCAAGAA 1115

AAAGACATAGTTCGCCGCTGTGAGGTTGTAGCCTGTGCCGGCGGCCGACCTGGTCCCGAT 1175

AAACACCCTGCAGTCCGGATCCTGCTGGAAAGCATCAATCGCCTTCTGCCGCTTCTTGGG 1235

CGAGTCACTGCCCACCAACGTCACGCACCCGACGCCAAGCTTGAGGCAGTGCTCCCGCAA 1295

CGTGGCCACGGATTCCTGATACTCGCAGAAGAGGATCACCTTGTCGTCGAC  1346
```

Fig. 3B

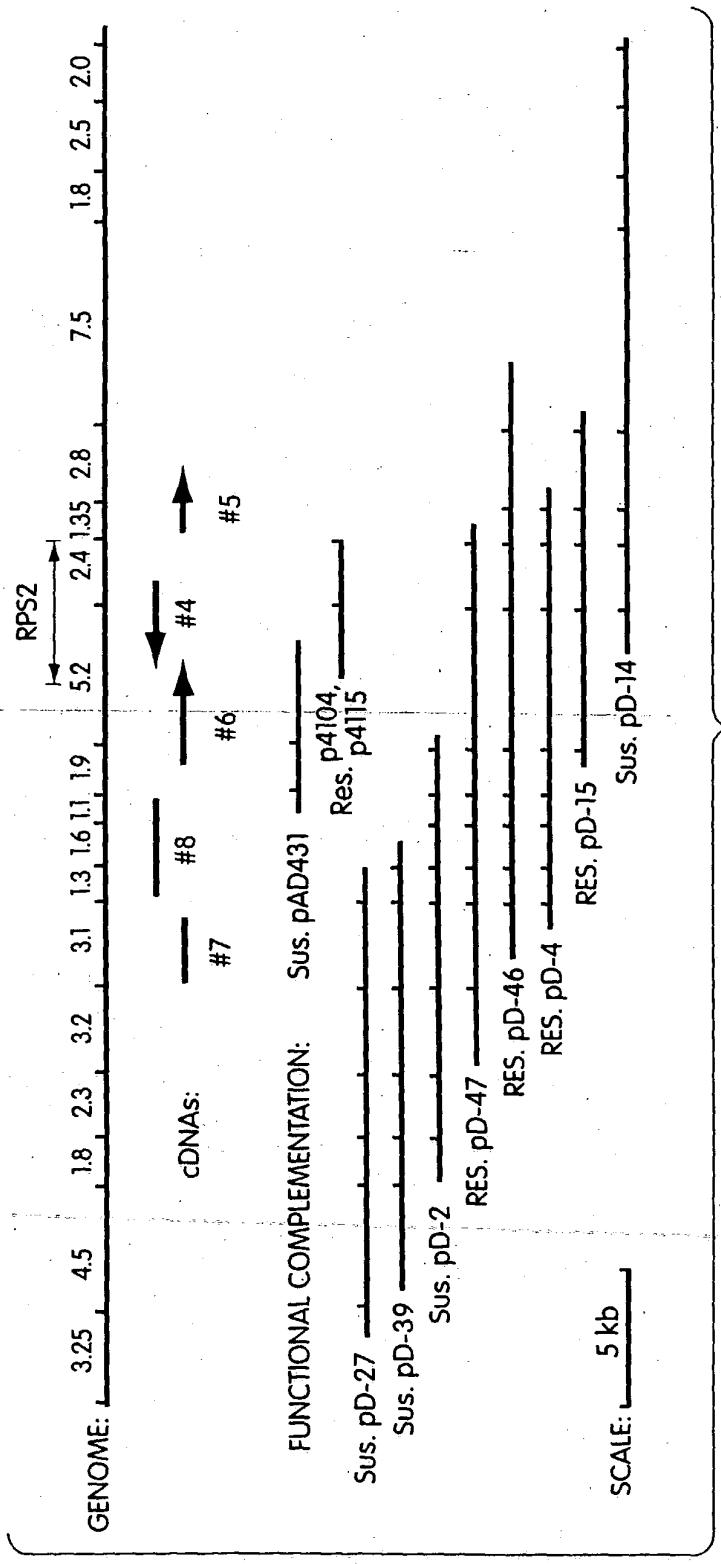

GENOME:
3.25 4.5 1.8 2.3 3.2 3.1 1.3 1.6 1.1 1.9 5.2 2.4 1.35 2.8 7.5 1.8 2.5 2.0
RPS2
cDNAs:
#7
#8
#6
#4
#5
FUNCTIONAL COMPLEMENTATION:
Sus. pAD431
Res. p4104, p4115
Sus. pD-27
Sus. pD-39
Sus. pD-2
RES. pD-47
RES. pD-46
RES. pD-4
RES. pD-15
Sus. pD-14
SCALE: 5 kb

Fig. 4

```
      1                                                        50
L6pro MSYLREVATA VALLLPFILL NKFWRPNSKD SIVNDDDDST SEVDAISDST
Nprot .......... .......... .......... .......... .........M
PrfP  .......... .......... .......... .......... ..........
rps2  .......... .......... .......... .......... ..........

      51              6                                       100
L6pro NPSGSFPSVE YEVFLSFRGP DTREQFTDFL YQSLRRYKIM TFRDDDELLK
Nprot ASSSSSSRWS YDVFLSFRGE DTRKTFTSHL YEVLNDKGIK TFQDDKRLEY
PrfP  .......... .......... .......... LRSKLDLIID LKHQIESVKE
rps2  .......... MDFISSLIVG CAQVLCESMN MAERRGHKTD LRQAITDLET

      101                                                     150
L6pro GKEIGPNLLR AIDQSKIYVP IISSGYADSK WCLMELAEIV RRQEEDPRRI
Nprot GATIPGELCK AIEESQFAIV VFSENYATSR WCLNELVKIM ECK.TRFKQT
PrfP  GLLCLRSFID HFSESYDEHD ........EA CGLIARVSVM AYKAE.....
rps2  AIGDLKAIRD DLTLRIQQDG LEGRSCSNRA REWLSAVQVT ETKTA.....

      151   7                                                 200
L6pro ILPIFYMVDP SDVRHQTGCY KKAFRKHANK F..DGQTIQN WKDALKKVGD
Nprot VIPIFYDVDP SHVRNQKESF AKAFEEHETK YKDDVEGIQR WRIALNEAAN
PrfP  .....YVIDS CLAYSHPLWY KVLW...... ..IS...... .EVLENIKLV
rps2  .....LLLVR FRRREQRTRM RRRY...... ..LSCFGCAD YKLCKKVSAI

      201                                                 8 250
L6pro LKGWHIGKND KQGAIADKVS ADIWSHISKE NLILE...TD ELVGIDDHIT
Nprot LKGSCDNRDK TDADCIRQIV DQISSKLCKI SLSY....LQ NIVGIDTHLE
PrfP  NKVVGETCER RNIEVTVHEV AKTTTYVAPS FSAYTQRANE EMEGFQDTID
rps2  LKSIGELRER SEAIKTDGGS IQVTCREIPI KSVVG..... .....NTTMM

      251                        1    -P-loop                300
L6pro AVLEKLSLDS ENVTMVGLYG MGGIGKTTTA KAVYNKI... ..SSC.FDCC
Nprot KIESLLEIGI NGVRIMGIWG MGGVGKTTIA RAIFDTLLGR MDSSYQFDGA
PrfP  ELKDKLLGGS PELDVISIVG MPGLGKTTLA KKIYNDPEVT ..SRFDVHAQ
rps2  EQVLEFLSEE EERGIIGVYG PGGVGKTTLM QSINNELITK ..G....HQY

      301                                                     350
L6pro CFIDNIRETQ EKDGVVVLQK KLVSEILRID ..SGSVGFNN DSGGRKTIKE
Nprot CFLKDIKE.. NKRGMHSLQN ALLSELLR.. ...EKANYNN EEDGKHQMAS
PrfP  CVVTQLYSWR EL.LLTILND VLEP...S.. ...DRNEKED GE.IADELRR
rps2  DVLIWVQMSR EF.GECTIQQ AVGA...RLG ..LSWDEKET GENRALKIYR

      351     2                                         3       400
L6pro RVSRFKILVV LDDVDEKFKF EDMLGSPKDF ISQ.SRFIIT SRSMRVLGTL
Nprot RLRSKKVLIV LDDIDNKDHY LEYLAGDLDW FGNGSRIIIT TRDKHLI...
PrfP  FLLTKRFLIL IDDVWDYKVW DNLCMCFSD. VSNRSRIILT TRLNDVAEYV
rps2  ALRQKRFLLL LDDVWEEIDL EKTGVPRPD. RENKCKVMFT TRSIALCNNM
```

Fig. 5A-1

```
      401                                                        450
L6pro NEN.QCKLYE VGSMSKPRSL ELFSKHAFKK NT....PPSY YETLANDVVD
Nprot .EK.NDIIYE VTALPDHESI QLFKQHAFGK EV....PNEN FEKLSLEVVN
PrfP  .KC.ESDPHH LRLFRDDESW TLLQKEVFQG E....SCPPE LEDVGFEISK
rps2  .GA.EYK.LR VEFLEKKHAW ELFCSKVWRK DLLESSSIRR LAEI...IVS

      451     4                                                  500
L6pro TTAGLPLTLK VIGSLLFKQE IAV..WEDTL EQL....RRT LNLDEVYDRL
Nprot YAKGLPLALK VWGSLLHNLR LTE..WKSAI EHM....KNN .SYSGIIDNV
PrfP  SCRGLPLSVV LVAGVLKQKK KTLDSWKVVE QSLS..SQRI GSLEESISII
rps2  KCGGLPLALI TLGGAMAH.R ETEEEWIHAS EVLTRFPAEM KGMNYVFALL

      501  5             9                                     550
L6pro KISYDALNPE .AKEIFLDIA CFFIGQ..NK EEPYYMWTDC NFYPASNIIF
Nprot KISYDGLEPK .QQEMFLDIA CFLRGE..EK DYILQILESC HIGAEYGLRI
PrfP  GFSYKNL.PH YLKPCFLYFG GFLQGKDIHD SKMTKLWVAE EFVQANN...
rps2  KFSYDNLESD LLRSCFLYCA LFPEEHSIEI EQLVEYWVGE GFLTSSHGVN

      551                                  10                    600
L6pro LIQRCMIQVG ........DD DEFKMHDQLR DMGREIVRRE DVLPWKRSRI
Nprot LIDKSLVFIS ........EY NQVQMHDLIQ DMGKYIVNFQ KD.PGERSRL
PrfP  .......... ........EK GQEDTRTRF. .LGRSYW... ..........
rps2  TIYKGYFLIG DLKAACLLET GDEKTQVKMH NVVRSFALWM ASEQGTYKEL

      601                                                        650
L6pro WSAEEGIDLL LNKKGSSKVK AISI.PWGVK YEFK.SECFL NLSELRYLHA
Nprot WLAKEVEEVM SNNTGTMAME AIWVSSYSST LRFS.NQAVK NMKRLRVFNM
PrfP  .......... .......... .......... .......... ..........
rps2  ILVEPSMGHT EAPKAENWRQ ALVISLLDNR IQTL.PEKLI CPKLTTLMLQ

      651                                                        700
L6pro REAMLTGDFN NLLPNLKWLE LPFYKHGEDD PPLTNYTMKN LII.VILEHS
Nprot GRSSTHYAID YLPNNLRCFV CTNYPW...E SFPSTFELKM LVH.LQLRH.
PrfP  .......... .......... .......... .......... ..........
rps2  QNSSLKKIPT GFFMHMPVLR VLDLSF.... TSITEIPLSI KYL.VELYHL

      701                                                        750
L6pro HITADDWGGW RHMMKMAERL KVVRLASNYS LYGRRVR... ..........
Nprot .......NSL RHLWTETKHL PSL..... ...RRID... ..........
PrfP  .......... .......... .......... .......... ..........
rps2  SMSGTKISVL PQELGNLRKL KHLDLQRTQF LQTIPRDAIC WLSKLEVLNL

      751                                                        800
L6pro .LSD.CWRFP KSIEVLSMTA IEMDEVDIGE LKKLKTLVLK FCPIQKISGG
Nprot .LSW.SKRLT RTPDFTGMPN LEY..VNLYQ CSNLEEVHHS LGCCSKVIGL
PrfP  .......... .......... .......... .......... ..........
rps2  YYSY.AGWEL QSFGEDEAEE LGFADLEYLE NLTTLGITVL SLETLKTLFE
```

Fig. 5A-2

```
      801                                                        850
L6pro TFGMLKGLRE L.CLEFNWGT NLREVVADIG QLSSLKVLKT TGAKEVEINE
Nprot YLNDCKSLKR F......... .......... .......... .PCVNVESLE
 PrfP .......... .......... .......... .......... ..........
 rps2 FGALHKHIQH L.HVEECNEL LYFNLPSLTN HGRNLRRLSI KSCHDLEYLV

      851                                                        900
L6pro FPLGLK.... ...ELSTSSR IPNLSQLLDL EVLKVYDCKD GFDMPPASPS
Nprot Y.LGLR.... ...SCDSLEK LPEIYGRMKP EI........ QIHMQGSGIR
 PrfP .......... .......... .......... .......... ..........
 rps2 TPADFENDWL PSLEVLTLHS LHNLTRVWGN SVSQDCLRNI RCINISHCNK

      901                                                        950
L6pro EDESSVWWKV SKLKSLQLEK TRINVNVVDD ASSGGHLPRY LLPTSLTYLK
Nprot ELPSSIFQYK THVTKLLL.. .WNMKNLVAL PSSICRL... ...KSLVSLS
 PrfP .......... .......... .......... .......... ..........
 rps2 LKNVSWVQKL PKLEVIELFD CREIEELISE HESPSVEDPT LFP.SLKTLR

      951                                                       1000
L6pro IYQCTEPTWL P.GIENLENL TSLEVNDIFQ TLGGDLDGLQ GLRSLEILRI
Nprot VSGCSKLESL PEEIGDLDNL RVFDASDTL. .......... ......ILRP
 PrfP .......... .......... .......... .......... ..........
 rps2 TRDLPELNSI LPSRFSFQKV ETLVITNCPR VKKLPFQERR TQMNLPTVYC

      1001                                                      1050
L6pro RKVNGLARIK GLKDLLCSST CKLRKFYITE CPDLIELLPC ELGGQTVVVP
Nprot P.........  .......SSI IRLNKLIILM FRGFKDGVHF EFPPVAEGLH
 PrfP .......... .......... .......... .......... ..........
 rps2 EEKWWKALEK DQPNEELCYL PRFVPN.... .......... ..........

      1051                                                      1100
L6pro SMAELTIRDC PRLEVGPMIR SLPKFPMLKK LDLAVANITK EEDLDAIGSL
Nprot SLEYLNL.SY CNLIDGGLPE EIGSLSSLKK LDLSRNNF.. EHLPSSIAQL
 PrfP .......... .......... .......... .......... ..........
 rps2 .......... .......... .......... .......... ..........

      1101                                                      1150
L6pro EELVSLELEL DDTSSGIERI VSSSKLQKLT TLVVKVPSLR EIEGLEELKS
Nprot GALQSLDLK. .......... .DCQRLTQLP ELPPELNELH .VDCHMALKF
 PrfP .......... .......... .......... .......... ..........
 rps2 .......... .......... .......... .......... ..........

      1151                                                      1200
L6pro LQDLYLEGCT SLGRLPLEKL KE......LD IGGCPDLTEL VQTVVAVPSL
Nprot IHDL.VTKRK KLHRVKLDDA HNDTMYNLFA YTMFQNISSM RHDISASDSL
 PrfP .......... .......... .......... .......... ..........
 rps2 .......... .......... .......... .......... ..........
```

Fig. 5A-3

```
      1201                                                      1250
L6pro RGLTIRDCPR LEVGPMIQSL PKFPMLNELT LSMVNITKED ELEVLGSLEE
Nprot .SLTV..... FTGQPYPEKI PSWFHHQGWD .SSVSVNLPE NWYIPDKFLG
 PrfP .......... .......... .......... .......... ..........
 rps2 .......... .......... .......... .......... ..........

      1251                                                      1300
L6pro LD.SLELTLD DTCSSIERIS FLSKLQKLTT LIVEVPSLRE IEGLAELKSL
Nprot FAVCYSRSLI DTTAHLIPVC .DDKMSRMTQ KLALSECDTE SSNYSEWD.I
 PrfP .......... .......... .......... .......... ..........
 rps2 .......... .......... .......... .......... ..........

      1301                                                      1350
L6pro RILYL..... .......... .EGCTSLERL WPDQQQLGSL KNLNVLDIQG
Nprot HFFFVPFAGL WDTSKANGKT PNDYGIIRLS FSGEEKMYGL RLLYKEGPEV
 PrfP .......... .......... .......... .......... ..........
 rps2 .......... .......... .......... .......... ..........

      1351                                        1387
L6pro CKSLSVDHLS ALKTTLPPRA RITWPDQPYR .......
Nprot NALLQMRENS NEPTEHSTGI RRTQYNNRTS FYELING
 PrfP .......... .......... .......... .......
 rps2 .......... .......... .......... .......
```

Fig. 5A-4

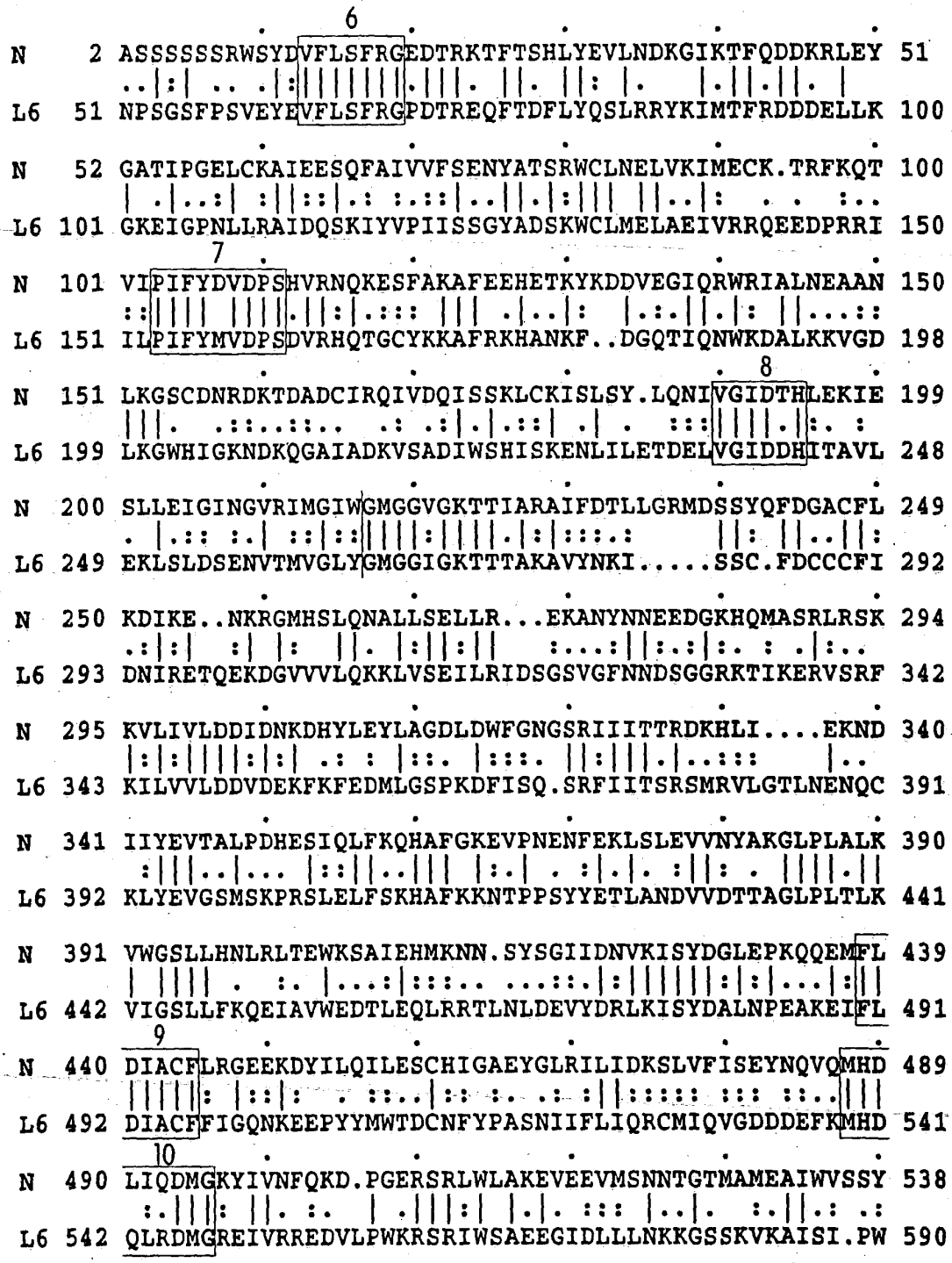

6
7
8
9
10
N 2 ASSSSSSRWSYDVFLSFRGEDTRKTFTSHLYEVLNDKGIKTFQDDKRLEY 51
L6 51 NPSGSFPSVEYEVFLSFRGPDTREQFTDFLYQSLRRYKIMTFRDDDELLK 100
N 52 GATIPGELCKAIEESQFAIVVFSENYATSRWCLNELVKIMECK.TRFKQT 100
L6 101 GKEIGPNLLRAIDQSKIYVPIISSGYADSKWCLMELAEIVRRQEEDPRRI 150
N 101 VIPIFYDVDPSHVRNQKESFAKAFEEHETKYKDDVEGIQRWRIALNEAAN 150
L6 151 ILPIFYMVDPSDVRHQTGCYKKAFRKHANKF..DGQTIQNWKDALKKVGD 198
N 151 LKGSCDNRDKTDADCIRQIVDQISSKLCKISLSY.LQNIVGIDTHLEKIE 199
L6 199 LKGWHIGKNDKQGAIADKVSADIWSHISKENLILETDELVGIDDHITAVL 248
N 200 SLLEIGINGVRIMGIWGMGGVGKTTIARAIFDTLLGRMDSSYQFDGACFL 249
L6 249 EKLSLDSENVTMVGLYGMGGIGKTTTAKAVYNKI.....SSC.FDCCCFI 292
N 250 KDIKE..NKRGMHSLQNALLSELLR...EKANYNNEEDGKHQMASRLRSK 294
L6 293 DNIRETQEKDGVVVLQKKLVSEILRIDSGSVGFNNDSGGRKTIKERVSRF 342
N 295 KVLIVLDDIDNKDHYLEYLAGDLDWFGNGSRIIITTRDKHLI....EKND 340
L6 343 KILVVLDDVDEKFKFEDMLGSPKDFISQ.SRFIITSRSMRVLGTLNENQC 391
N 341 IIYEVTALPDHESIQLFKQHAFGKEVPNENFEKLSLEVVNYAKGLPLALK 390
L6 392 KLYEVGSMSKPRSLELFSKHAFKKNTPPSYYETLANDVVDTTAGLPLTLK 441
N 391 VWGSLLHNLRLTEWKSAIEHMKNN.SYSGIIDNVKISYDGLEPKQQEMFL 439
L6 442 VIGSLLFKQEIAVWEDTLEQLRRTLNLDEVYDRLKISYDALNPEAKEIFL 491
N 440 DIACFLRGEEKDYILQILESCHIGAEYGLRILIDKSLVFISEYNQVQMHD 489
L6 492 DIACFFIGQNKEEPYYMWTDCNFYPASNIIFLIQRCMIQVGDDDEFKMHD 541
N 490 LIQDMGKYIVNFQKD.PGERSRLWLAKEVEEVMSNNTGTMAMEAIWVSSY 538
L6 542 QLRDMGREIVRREDVLPWKRSRIWSAEEGIDLLLNKKGSSKVKAISI.PW 590

Fig. 5B-1

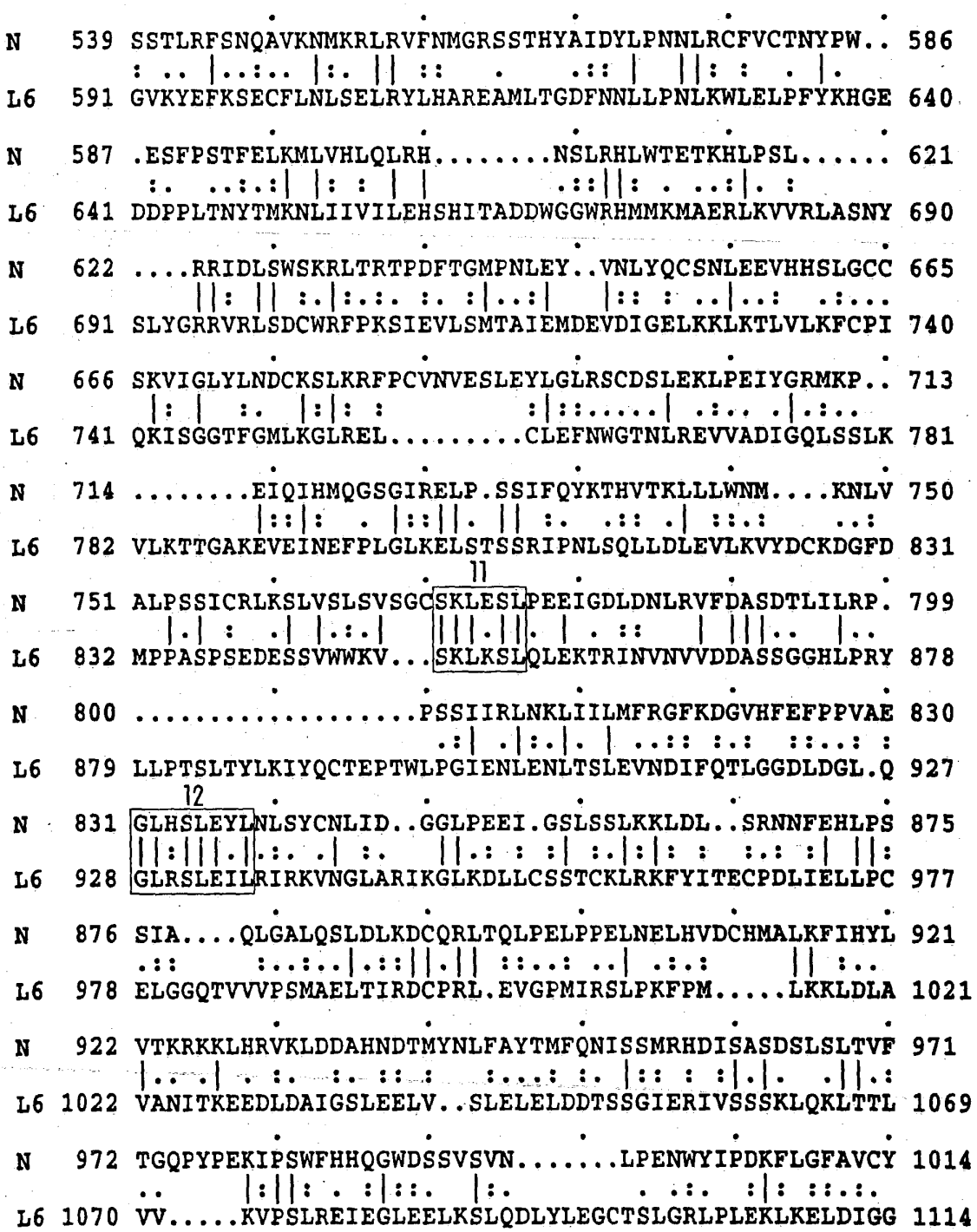


Fig. 5B-2

```
N   1015 SRSLIDTTAHLIPVCDDK..........MSRMTQKLA....LSECDTES 1049
         :..|.: .. :::|.. :           ::.|.|.|:    |.| . .
L6  1115 CPDLTELVQTVVAVPSLRGLTIRDCPRLEVGPMIQSLPKFPMLNELTLSM 1164

N   1050 SNYSEWDIHFFFVPFAGLWDTSKANGKTPNDYGIIRLSFSGEEKMYGLRL 1099
          |... |   .:..:.:| . . . :.|....: |.: :|  :|: .| :
L6  1165 VNITKEDELEVLGSLEELDSLELTLDDTCSSIERISF.LSKLQKLTTLIV 1213

N   1100 LYKEGPEVNALLQMRENSNEPTEHSTGIRRTQYNNRTSFYELIN 1143
           .. .|:::| :::. .    | :|:: |  : :.  : .| |
L6  1214 EVPSLREIEGLAELKSLRILYLEGCTSLER.LWPDQQQLGSLKN 1256
```

Fig. 5B-3

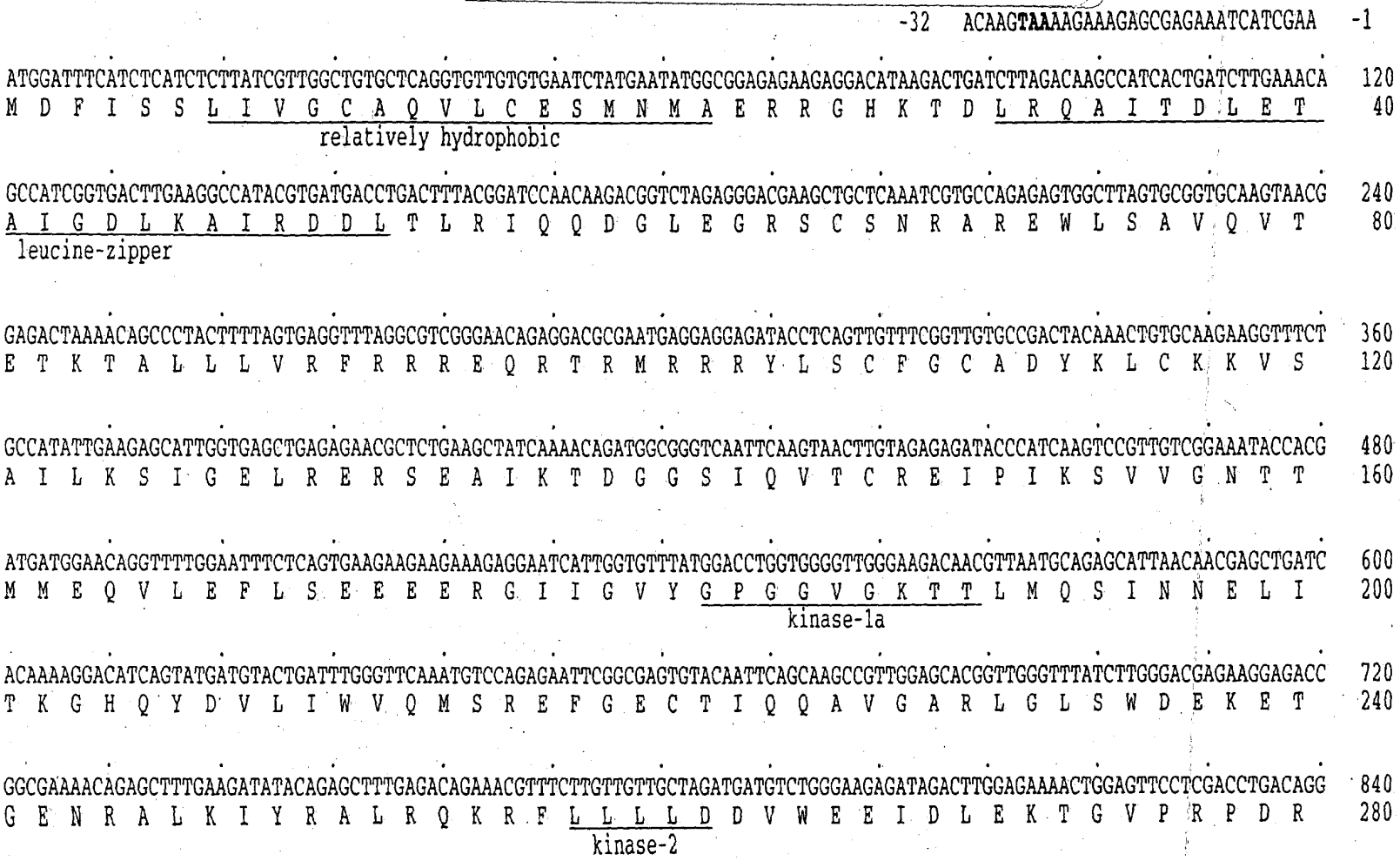

-32 ACAAGTAAAAGAAAGAGCGAGAAATCATCGAA -1
ATGGATTTCATCTCATCTCTTATCGTTGGCTGTGCTCAGGTGTTGTGTGAATCTATGAATATGGCGGAGAGAAGAGGACATAAGACTGATCTTAGACAAGCCATCACTGATCTTGAAACA 120
M D F I S S L I V G C A Q V L C E S M N M A E R R G H K T D L R Q A I T D L E T 40
relatively hydrophobic
GCCATCGGTGACTTGAAGGCCATACGTGATGACCTGACTTTACGGATCCAACAAGACGGTCTAGAGGGACGAAGCTGCTCAAATCGTGCCAGAGAGTGGCTTAGTGCGGTGCAAGTAACG 240
A I G D L K A I R D D L T L R I Q Q D G L E G R S C S N R A R E W L S A V Q V T 80
leucine-zipper
GAGACTAAAACAGCCCTACTTTTAGTGAGGTTTAGGCGTCGGGAACAGAGGACGCGAATGAGGAGGAGATACCTCAGTTGTTTCGGTTGTGCCGACTACAAACTGTGCAAGAAGGTTTCT 360
E T K T A L L L V R F R R R E Q R T R M R R R Y L S C F G C A D Y K L C K K V S 120
GCCATATTGAAGAGCATTGGTGAGCTGAGAGAACGCTCTGAAGCTATCAAAACAGATGGCGGGTCAATTCAAGTAACTTGTAGAGAGATACCCATCAAGTCCGTTGTCGGAAATACCACG 480
A I L K S I G E L R E R S E A I K T D G G S I Q V T C R E I P I K S V V G N T T 160
ATGATGGAACAGGTTTTGGAATTTCTCAGTGAAGAAGAAGAAAGAGGAATCATTGGTGTTTATGGACCTGGTGGGGTTGGGAAGACAACGTTAATGCAGAGCATTAACAACGAGCTGATC 600
M M E Q V L E F L S E E E E R G I I G V Y G P G G V G K T T L M Q S I N N E L I 200
kinase-1a
ACAAAAGGACATCAGTATGATGTACTGATTTGGGTTCAAATGTCCAGAGAATTCGGCGAGTGTACAATTCAGCAAGCCGTTGGAGCACGGTTGGGTTTATCTTGGGACGAGAAGGAGACC 720
T K G H Q Y D V L I W V Q M S R E F G E C T I Q Q A V G A R L G L S W D E K E T 240
GGCGAAAACAGAGCTTTGAAGATATACAGAGCTTTGAGACAGAAACGTTTCTTGTTGTTGCTAGATGATGTCTGGGAAGAGATAGACTTGGAGAAAACTGGAGTTCCTCGACCTGACAGG 840
G E N R A L K I Y R A L R Q K R F L L L L D D V W E E I D L E K T G V P R P D R 280
kinase-2

Fig. 6A

```
GAAAACAAATGCAAGGTGATGTTCACGACACGGTCTATAGCATTATGCAACAATATGGGTGCGGAATACAAGTTGAGAGTGGAGTTTCTGGAGAAGAAACACGCGTGGGAGCTGTTCTGT  960
E  N  K  C  K  V  M  F  T  T  R  S  I  A  L  C  N  N  M  G  A  E  Y  K  L  R  V  E  F  L  E  K  K  H  A  W  E  L  F  C   320

AGTAAGGTATGGAGAAAAGATCTTTTAGAGTCATCATCAATTCGCCGGCTCGCGGAGATTATAGTGAGTAAATGTGGAGGATTGCCACTAGCGTTGATCACTTTAGGAGGAGCCATGGCT  1080
S  K  V  W  R  K  D  L  L  E  S  S  S  I  R  R  L  A  E  I  I  V  S  K  C  G  G  L  P  L  A  L  I  T  L  G  G  A  M  A   360
                           kinase-3a                                                    membrane integarated

CATAGAGAGACAGAAGAAGAGTGGATCCATGCTAGTGAAGTTCTGACTAGATTTCCAGCAGAGATGAAGGGTATGAACTATGTATTTGCCCTTTTGAAATTCAGCTACGACAACCTCGAG  1200
H  R  E  T  E  E  E  W  I  H  A  S  E  V  L  T  R  F  P  A  E  M  K  G  M  N  Y  V  F  A  L  L  K  F  S  Y  D  N  L  E   400

AGTGATCTGCTTCGGTCTTGTTTCTTGTACTGCGCTTTATTCCCAGAAGAACATTCTATAGAGATCGAGCAGCTTGTTGAGTACTGGGTCGGCGAAGGGTTTCTCACCAGCTCCCATGGC  1320
S  D  L  L  R  S  C  F  L  Y  C  A  L  F  P  E  E  H  S  I  E  I  E  Q  L  V  E  Y  W  V  G  E  G  F  L  T  S  S  H  G   440

GTTAACACCATTTACAAGGGATATTTTCTCATTGGGGATCTGAAAGCGGCATGTTTGTTGGAAACCGGAGATGAGAAAACACAGGTGAAGATGCATAATGTGGTCAGAAGCTTTGCATTG  1440
V  N  T  I  Y  K  G  Y  F  L  I  G  D  L  K  A  A  C  L  L  E  T  G  D  E  K  T  Q  V  K  M  H  N  V  V  R  S  F  A  L   480

TGGATGGCATCTGAACAGGGGACTTATAAGGAGCTGATCCTAGTTGAGCCTAGCATGGGACATACTGAAGCTCCTAAAGCAGAAAACTGGCGACAAGCGTTGGTGATCTCATTGTTAGAT  1560
W  M  A  S  E  Q  G  T  Y  K  E  L  I  L  V  E  P  S  M  G  H  T  E  A  P  K  A  E  N  W  R  Q  A  L  V  I  S  L  L  D   520

AACAGAATCCAGACCTTGCCTGAAAAACTCATATGCCCGAAACTGACAACACTGATGCTCCAACAGAACAGCTCTTTGAAGAAGATTCCAACAGGGTTTTTCATGCATATGCCTGTTCTC  1680
N  R  I  Q  T  L  P  E  K  L  I  C  P  K  L  T  T  L  M  L  Q  Q  N  S  S  L  K  K  I  P  T  G  F  F  M  H  M  P  V  L   560

AGAGTCTTGGACTTGTCGTTCACAAGTATCACTGAGATTCCGTTGTCTATCAAGTATTTGGTGGAGTTGTATCATCTGTCTATGTCAGGAACAAAGATAAGTGTATTGCCACAGGAGCTT  1800
R  V  L  D  L  S  F  T  S  I  T  E  I  P  L  S  I  K  Y  L  V  E  L  Y  H  L  S  M  S  G  T  K  I  S  V  L  P  Q  E  L   600
```

Fig. 6B

```
GGGAATCTTAGAAAACTGAAGCATCTGGACCTACAAAGAACTCAGTTTCTTCAGACGATCCCACGAGATGCCATATGTTGGCTGAGCAAGCTCGAGGTTCTGAACTTGTACTACAGTTAC 1920
G  N  L  R  K  L  K  H  L  D  L  Q  R  T  Q  F  L  Q  T  I  P  R  D  A  I  C  W  L  S  K  L  E  V  L  N  L  Y  Y  S  Y   640

GCCGGTTGGGAACTGCAGAGCTTTGGAGAAGATGAAGCAGAAGAACTCGGATTCGCTGACTTGGAATACTTGGAAAACCTAACCACACTCGGTATCACTGTTCTCTCATTGGAGACCCTA 2040
A  G  W  E  L  Q  S  F  G  E  D  E  A  E  E  L  G  E  A  D  L  E  Y  L  E  N  L  T  T  L  G  I  T  V  L  S  L  E  T  L   680

AAAACTCTCTTCGAGTTCGGTGCTTTGCATAAACATATACAGCATCTCCACGTTGAAGAGTGCAATGAACTCCTCTACTTCAATCTCCCATCACTCACTAACCATGGCAGGAACCTGAGA 2160
K  T  L  F  E  F  G  A  L  H  K  H  I  Q  H  L  H  V  E  E  C  N  E  L  L  Y  F  N  L  P  S  L  T  N  H  G  R  N  L  R   720

AGACTTAGCATTAAAAGTTGCCATGACTTGGAGTACCTGGTCACACCCGCAGATTTTGAAAATGATTGGCTTCCGAGTCTAGAGGTTCTGACGTTACACAGCCTTCACAACTTAACCAGA 2280
R  L  S  I  K  S  C  H  D  L  E  Y  L  V  T  P  A  D  F  E  N  D  W  L  P  S  L  E  V  L  T  L  H  S  L  H  N  L  T  R   760

GTGTGGGGAAATTCTGTAAGCCAAGATTGTCTGCGGAATATCCGTTGCATAAACATTTCACACTGCAACAAGCTGAAGAATGTCTCATGGGTTCAGAAACTCCCAAAGCTAGAGGTGATT 2400
V  W  G  N  S  V  S  Q  D  C  L  R  N  I  R  C  I  N  I  S  H  C  N  K  L  K  N  V  S  W  V  Q  K  L  P  K  L  E  V  I   800

GAACTGTTCGACTGCAGAGAGATAGAGGAATTGATAAGCGAACACGAGAGTCCATCCGTCGAAGATCCAACATTGTTCCCAAGCCTGAAGACCTTGAGAACTAGGGATCTGCCAGAACTA 2520
E  L  F  D  C  R  E  I  E  E  L  I  S  E  H  E  S  P  S  V  E  D  P  T  L  F  P  S  L  K  T  L  R  T  R  D  L  P  E  L   840

AACAGCATCCTCCCATCTCGATTTTCATTCCAAAAAGTTGAAACATTAGTCATCACAAATTGCCCCAGAGTTAAGAAACTGCCGTTTCAGGAGAGGAGGACCCAGATGAACTTGCCAACA 2640
N  S  I  L  P  S  R  F  S  F  Q  K  V  E  T  L  V  I  T  N  C  P  R  V  K  K  L  P  F  Q  E  R  R  T  Q  M  N  L  P  T   880
```

Fig. 6C

```
GTTTATTGTGAGGAGAAATGGTGGAAAGCACTGGAAAAAGATCAACCAAACGAAGAGCTTTGTTATTTACCGCGCTTTGTTCCAAATTGATATAAGAGCTAAGAGCACTCTGTACAAATA  2760
V  Y  C  E  E  K  W  W  K  A  L  E  K  D  Q  P  N  E  E  L  C  Y  L  P  R  F  V  P  N  *                                  909

TGTCCATTCATAAGTAGCAGGAAGCCAGGAAGGTTGTTCCAGTGAAGTCATCAACTTTCCACTAGACCACAAAACTAGAGATTATGTAATCATAAAAACCAAACTATCCGCGATCAAATA  2880

GATCTCACGACTATGAGGACGAAGACTCACCGAGTATCGTCGATATAGAAACTCCAAGCTCCAGTTCCGATCAGTGAAGACGAACAAGTTTATCAGATCTCTGCAACAATTCTGGGAATC  3000

GTCACCTCAGATTAGACCTCCAGTAAGAAGTGAGAAAGCATGGACGACGACTGTGAAGAATTGAGCTAATGAGCTGAACCGGATCCGGTGAAATTGCAGAACCGGATCGGAGAAGAAGAA  3120

TTTTGCATTTGTGCATCTTTATTTTTAATTGTTACGTTTGAGCCCCAATAATCATAGATATTGTAGTGAAGACCAAATTTCATGGTGGATCAATCAAATTGTATTTTCAAATTTTCGTAG  3240

TGTAATAACGGAAAAAGGAATAAAAAGGTCACTGAGT (A)n
```

Fig. 6D

```
consensus   PXXaXX LXXLXXLXaXXXX aXXa

          505   PKAENW RQALVISLLD NR IQTL
          527   PEKLIC PK LTTLMLQQNSSLKKI
          550   PTGFFMHMPVLRVLDLSFTS ITEI
          574   PLSIKY LVELYHLSMSGTK ISVL
          597   PQELGN LRKLKHLDLQRTQFLQTI
          621   PRDAICWLSKLEVLNLYYSYAGWEL  QSFGEDEAEELG
          658   FADLEY LENLTTLGITVLS LETL  KT
          683   LFEFGALHKHIQHLHVEECNELLYF  NL
          710   P SLTNHGRNLRRLSIKSCHDLEYL  VT
          736  PADFENDWLPSLEVLTLHSLHNLTRV  WGN
          765   SVSQDC LRNIRCINISHCNKLKNV  SWVQKL
          795   PK LEV IELFDCREIEELISEHES  PSVED
          823   PT LFPSLKTLRTRDLPELNSI  L
          845   PSRFS  FQKVETLVITNCPRVKKL
```

Fig. 7

Leucine zipper

MDFISSLIVG CAQVLCESMN MAERRGHKTD LRQAITDLET AIGDLKAIRD DLTLRIQQDG 120
LEGRSCSNRA REWLSAVQVT ETKTALLLVR FRRREQRTRM RRRYLSCFGC ADYKLCKKVS 180
AILKSIGELR ERSEAIKTDG GSIQVTCREI PIKSVVGNTT MMEQVLEFLS EEEERGIIGV 240
P loop
YGPGGVGKTT LMQSINNELI TKGHQYDVLI WVQMSREFGE CTIQQAVGAR LGLSWDEKET 300
GENRALKIYR ALRQKRFLLL LDDVWEEIDL EKTGVPRPDR ENKCKVMFTT RSIALCNNMG 360
Membrane-spanning
AEYKLRVEFL EKKHAWELFC SKVWRKDLLE SSSIRRLAEI IVSKCGGLPL ALITLGGAMA 420
HRETEEEWIH ASEVLTRFPA EMKGMNYVFA LLKFSYDNLE SDLLRSCFLY CALFPEEHSI 480
EIEQLVEYWV GEGFLTSSHG VNTIYKGYFL IGDLKAACLL ETGDEKTQVK MHNVVRSFAL 540
WMASEQGTYK ELILVEPSMG HTEAPKAENW RQALVISLLD NRIQTLPEKL ICPKLTTLML 600
◄—— Leucine-rich repeats ——►
QQNSSLKKIP TGFFMHMPVL RVLDLSFTSI TEIPLSIKYL VELYHLSMSG TKISVLPQEL 660
GNLRKLKHLD LQRTQFLQTI PRDAICWLSK LEVLNLYYSY AGWELQSFGE DEAEELGFAD 720
LEYLENLTTL GITVLSLETL KTLFEFGALH KHIQHLHVEE CNELLYFNLP SLTNHGRNLR 780
RLSIKSCHDL EYLVTPADFE NDWLPSLEVL TLHSLHNLTR VWGNSVSQDC LRNIRCINIS 840
(end Leucine-rich repeats)
HCNKLKNVSW VQKLPKLEVI ELFDCREIEE LISEHESPSV EDPTLFPSLK TLRTRDLPEL 900
NSILPSRFSF QKVETLVITN CPRVKKLPFQ ERRTQMNLPT VYCEEKWWKA LEKDQPNEEL
909
CYLPRFVPN

Fig. 8

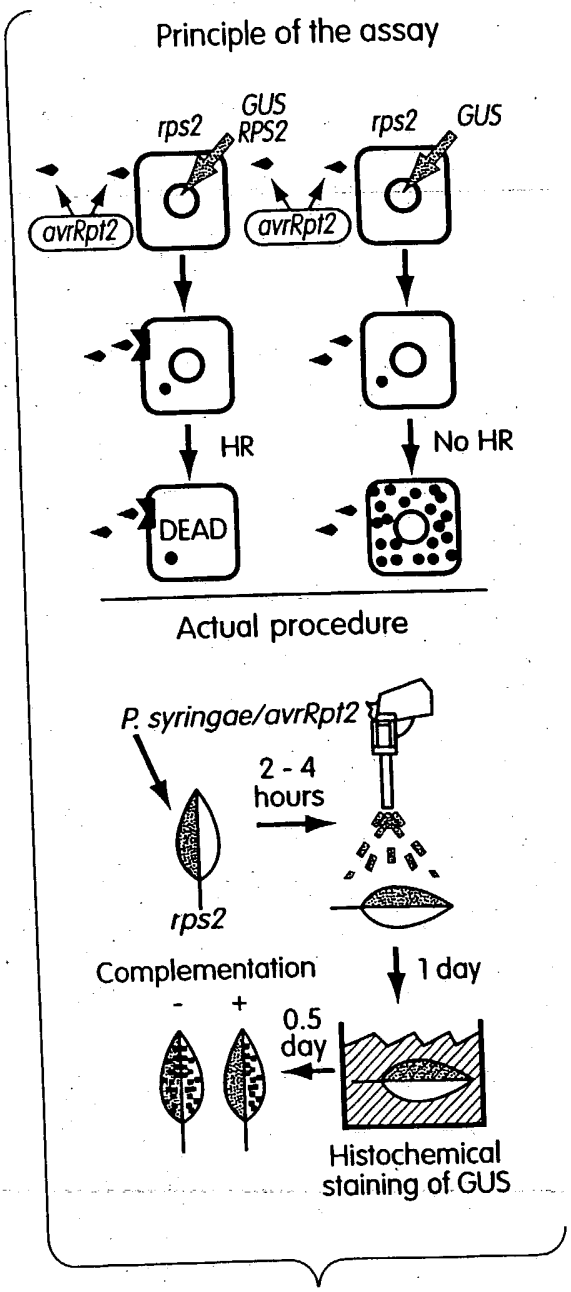

Principle of the assay
rps2
GUS
RPS2
avrRpt2
rps2
GUS
avrRpt2
HR
No HR
DEAD
Actual procedure
P. syringae/avrRpt2
2 - 4
hours
rps2
Complementation
1 day
-
+
0.5
day
Histochemical
staining of GUS

Fig. 9

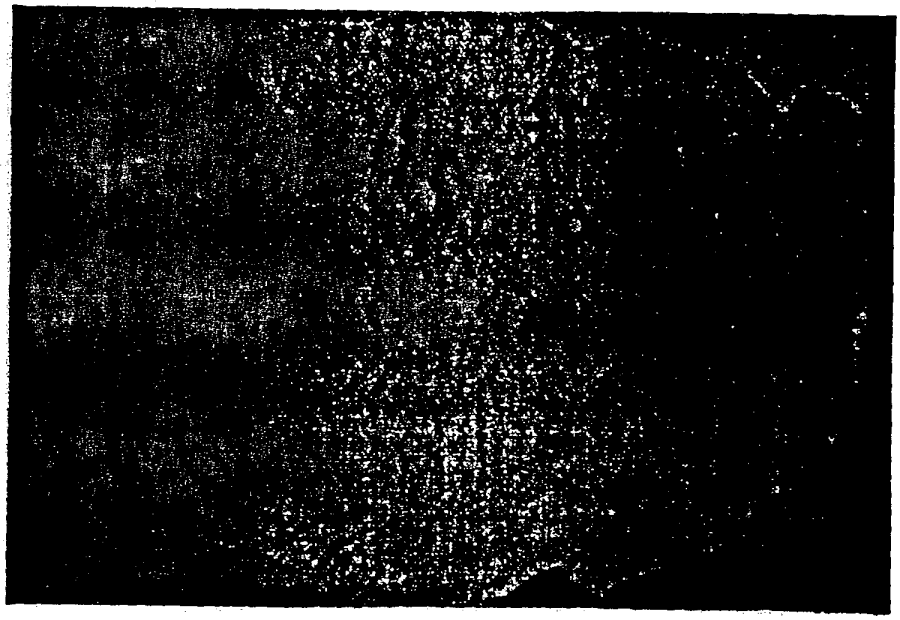


Fig. 10B

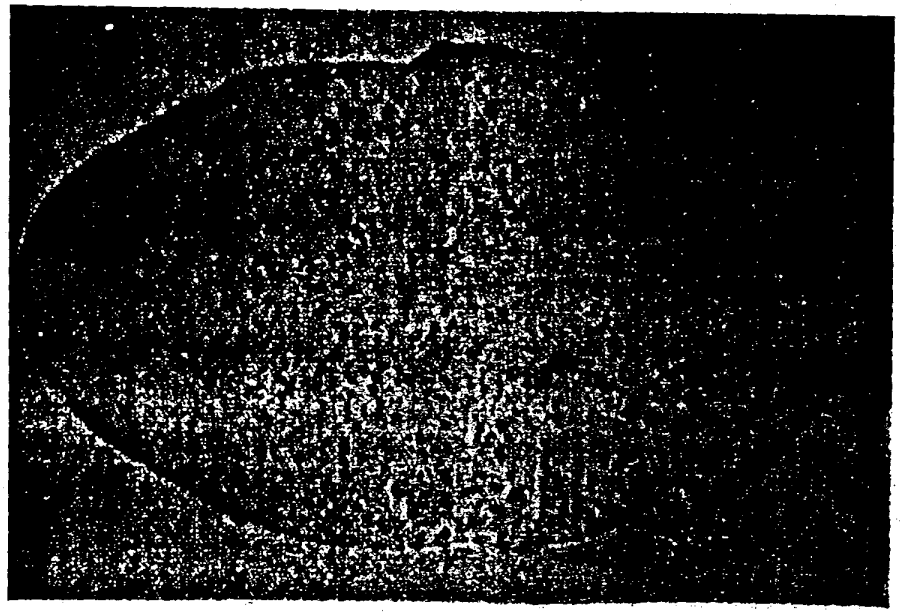


Fig. 10A

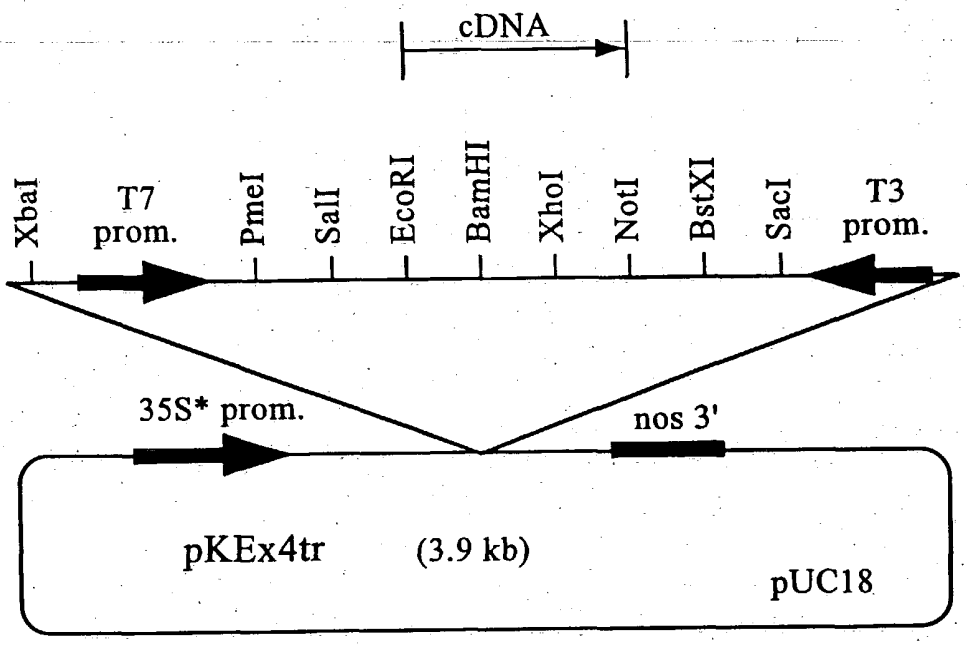

cDNA
XbaI
T7 prom.
PmeI
SalI
EcoRI
BamHI
XhoI
NotI
BstXI
SacI
T3 prom.
35S* prom.
nos 3'
pKEx4tr
(3.9 kb)
pUC18

Fig. 11

```
      |  10      |  20      |  30      |  40      |  50      |  60
    1 aagctttaca gattggatga tctcttaatg catgctgaag tgactgcaaa aaggttagca  60
   61 atattcagtg gttctcgtta tgaatatttc atgaacggaa gcagcactga gaaaatgagg  120
  121 cccttgttat ctgattttct gcaagagatt gagtctgtca aggtagagtt cagaaatgtt  180
  181 tgcttgcaag ttctggatat atcacctttt tccctgacag atggagaagg ccttgttaat  240
  241 ttcttattaa aaaaccaggc caaggtgccg aatgatgatg ctgtttcttc tgatggaagt  300
  301 ttagaggatg caagcagcac tgagaaaatg ggacttccat ctgattttct ccgagagatt  360
  361 gagtctgttg agataaagga ggccagaaaa ttatatgatc aagttttgga tgcaacacat  420
  421 tgtgagacga gtaagcacga tggaaaaagc tttatcaaca ttatgttaac ccaacaggac  480
  481 aaggtgctgg actatgatgc tggttcagtg tcttatcttc ttaaccaaat ctcagtagtt  540
  541 aaagacaaaa tattgcacat tggctcttta cttgtagata ttgtacagta ccggaatatg  600
  601 catatagaac ttacagatct cgctgaacgt gttcaagata aaaactacat tcgtttcttc  660
  661 tctgtcaagg gttatattcc tgcttggtat tacacactat atctctctga tgtcaagcaa  720
  721 ttgcttaagt ttgttgaggc agaggtaaag attatttgtc tgaaagtacc agattcttca  780
  781 agttatagct tccctaagac aaatggatta ggatatctca attgcttttt aggcaaattg  840
  841 gaggagcttt tacgttctaa gctcgatttg ataatcgact taaaacatca gattgaatca  900
  901 gtcaaggagg gcttattgtg cctaagatca ttcattgatc atttttcaga aagctatgtt  960
  961 gagcatgatg aagcttgtgg tcttatagca agagtttctg taatggcata caaggctgag 1020
 1021 tatgtcattg actcatgctt ggcctattct catccactct ggtacaaagt tctttggatt 1080
 1081 tctgaagttc ttgagaatat taagcttgta aataaagttg ttggggagac atgtgaaaga 1140
 1141 aggaacactg aagttactgt gcatgaagtt gcaaagacta ccactaatgt agcaccatct 1200
 1201 ttttcagctt atactcaaag agcaaacgaa gaaatggagg gttttcagga tacaatagat 1260
 1261 gaattaaagg ataaactact tggaggatca cctgagcttg atgtcatctc aatcgttggc 1320
 1321 atgccaggat tgggcaagac tacactagca aagaagattt acaatgatcc agaagtcacc 1380
 1381 tctcgcttcg atgtccatgc tcaatgtgtt gtgactcaat tatattcatg gagagagttg 1440
 1441 ttgctcacca ttttgaatga tgtgcttgag ccttctgatc gcaatgaaaa agaagatgga 1500
 1501 gaaatagctg atgatctacg ccgatttttg ttgaccaaga gattcttgat tctcattgat 1560
 1561 gatgtgtggg actataaagt gtgggacaat ctatgtatgt gcttcagtga tgtttcaaat 1620
 1621 aggagtagaa ttatcctaac aacccgcttg aatgatgtcg ccgaatatgt caaatgtgaa 1680
 1681 agtgatcccc atcatcttcg tttattcaga gatgacgaga gttggacatt attacagaaa 1740
 1741 gaagtctttc aaggagagag ctgtccacct gaacttgaag atgtgggatt tgaaatatca 1800
 1801 aaaagttgta gagggttgcc tctctcagtt gtgttagtag ctggtgttct gaaacagaaa 1860
 1861 aagaagacac tagattcatg gaaagtagta gaacaaagtc taagttccca gaggattggc 1920
 1921 agcttggaag agagcatatc tataattgga ttcagttaca agaatttacc acactatctt 1980
 1981 aagccttgtt ttctctattt tggaggattt ttgcagggaa aggatattca tgactcaaaa 2040
 2041 atgaccaagt tgtgggtagc tgaagagttt gtacaagcaa acaacgaaaa aggacaagaa 2100
 2101 gatacccgca caaggtttct tggacgatct tattggtagg aatctggtga tggccatgga 2160
 2161 gaagagacct aatgccaagg tgaaaacgtg ccgcattcat gatttgttgc ataaattctg 2220
 2221 catggaaaag gccaaacaag aggatttcct tctccagatc aataggtaaa aaaaactgta 2280
 2281 ttaattttac attacaaaaa aaaagaactg tattaatttt actgtattat gtttatgcca 2340
 2341 actctcattt ccatgtgttc tcttttattc aattcagtgg agaaggtgta tttcctgaac 2400
 2401 gattggaaga ataccgattg ttcgttcatt cttaccaaga tgaaattgat ctgtggcgcc 2460
 2461 catctcgctc taatgtccgc tctttactat tcaatgcaat tgatccagat aacttgttat 2520
 2521 ggccgcgtga tatctccttc atttttgaga gcttcaagct tgttaaagtg ttggatttgg 2580
```

Fig. 12A

```
2581 aatcattcaa cattggtggt acttttccca ttgaaacaca atatctaatt cagatgaagt 2640
2641 actttgcggc ccaaactgat gcaaattcaa ttccttcatc tatagctaag cttgaaaatc 2700
2701 ttgagacttt tgtcgtaaga ggattgggag gagagatgat attaccttgt tcacttctga 2760
2761 agatggtgaa attgaggcat atacatgtaa atgatcgggt ttcttttggt ttgcgtgaga 2820
2821 acatggatgt tttaactggt aactcacaat aacctaattt ggaaaccttt tctactccgc 2880
2881 gtctctttta tggtaaagac gcagagaaga ttttgaggaa gatgccaaaa ttgagaaaat 2940
2941 tgagttgcat attttcaggg acatttggtt attcaaggaa attgaagggt aggtgtgttc 3000
3001 gttttcccag attagatttt ctaagtcacc ttgagtccct caagctggtt tcgaacagct 3060
3061 atccagccaa acttcctcac aagttcaatt tcccctcgca actaagggaa ctgactttat 3120
3121 caaagttccg tctaccttgg acccaaattt cgatcattgc agaactgccc aacttggtga 3180
3181 ttcttaagtt attgctcaga gcctttgaag gggatcactg ggaagtgaaa gattcagagt 3240
3241 tcctagaact caaatactta aaactggaca acctcaaagt tgtacaatgg tccatctctg 3300
3301 atgatgcttt tcctaagctt gaacatttgg ttttaacgaa atgtaagcat cttgagaaaa 3360
3361 tcccttctcg ttttgaagat gctgtttgtc taaatagagt tgaggtgaac tggtgcaact 3420
3421 ggaatgttgc caattcagcc caagatattc aaactatgca acatgaagtt atagcaaatg 3480
3481 attcattcac agttactata cagcctccag attggtctaa agaacagccc cttgactctt 3540
3541 agcaaaggtt tgttcttgct gtgttcatcc aagtgcattt aacatttatt cattttgttt 3600
3601 tacaccagaa catgtttatt ttgctagtat tacttgatac attaaaagaa atcgaactca 3660
3661 tatttctgct acagtcttaa cttttcttgg gcttacttga ggtctagatt agatcaatgg 3720
3721 ttcatgtaat ttttaattca ctgtttcatt caactgtctt atgatagttg tgaaatgaca 3780
3781 atattgttat ccctagccaa atttattatg ttcaaatgaa aactgatgtc acaactactt 3840
3841 ttttgtgaaa tgtttttgaa ttttttgcta taaaattgac gaattgacag cttctatatt 3900
3901 tgtcagctaa actctttgtc accagaagtg tatttagaat tactgtggtt ttatgaaaga 3960
3961 gttctgtaga attttatgct tttgcagaat atagtttaaa acaacaacac ttctctgttt 4020
4021 cagagatagc agaagctaaa gttcaaggca ttttgtttat ttctagaaca agtggagttc 4080
4081 ttatgttgaa ttcttgaaaa gaagaagaat caggagcagg taaagttatc tctttttatg 4140
4141 tttttcttct tttagatgtt atttcttcat cttgaacgtg aacaccgctg aaagcatttt 4200
4201 aataaaaccg gagagaaaaa taagatcttt ttatataaag cattatcatg taaatatgcc 4260
4261 taaatccata tggtacaact gtttgacaaa atgatagaga ggggagtttt atagtataag 4320
4321 taaaacagga ttgagaaaaa aatccttgca cgattttcaa tttctggcca catcacaatg 4380
4381 tgtgtcaaag ttccctctt taagtggaac aagcaatcag aaaagctcat tcttatcggt 4440
4441 gacataccaa taccagctga ctgtctcatc ttggttaact tagccttgct tacttagact 4500
4501 attagattag ttactaatga actggtaaat tggaaccaaa tgtagttagc ttgatgagct 4560
4561 ggtagacatg tatatatgaa gatacacgcg taactttagt cgatggttaa tttttcattt 4620
4621 ttgatttttt ttcttcacag agtatatatg aacttggcct aaaagttttg cttcactaat 4680
4681 ttaactatta ccgtggatga aacaagcatg gcaacatttt caacaactat cactcaagca 4740
4741 atgtaaaaaa tggaggttct acgagcggta catgtaagag ttttgtgcac acaagaggtt 4800
4801 ctgagacttg aaccatccat gtccaaggca gttgagatgc tagtaaagaa agaagaagat 4860
4861 gagcctgcac taattaatct ccctgtatga atgagagaat gagaaaaaga tggagcttca 4920
4921 tgaaccaaaa gttacctttt ttttttcttc ttaatggcat tactttgaag cacatgtttg 4980
4981 ttagttgtaa attgtaatgg tgaagtgttt gtaaatatag ggagtgatat ttgaaagaat 5040
5041 ggttgtgtta tctttacaaa ccggaatcat ttctgtataa ttttcttctg taatttttgg 5100
5101 tttcggttta ttcattactc atttcagtaa gctt                          5134
         |  10     |  20     |  30     |  40     |  50     |  60
```

Fig. 12B